Treatment Outcomes of Pre-surgical Infant Orthopedics in Patients with Non-Syndromic Cleft Lip and/or Palate: A Systematic Review and Meta-analysis of Randomized Controlled Trials

Hamid Reza Hosseini
DDS, Ajman University of Science and Technology, 2009
PGDip, Queen Mary University, 2012

Presented to the Hamdan Bin Mohammed College of Dental Medicine
Mohammed Bin Rashid University of Medicine and Health Sciences
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Orthodontics
2016

ABSTRACT

TREATMENT OUTCOMES OF PRE-SURGICAL INFANT ORTHOPEDICS IN PATIENTS WITH NON-SYNDROMIC CLEFT LIP AND/OR PALATE: A SYSTEMATIC REVIEW AND META-ANALYSIS OF RANDOMIZED CONTROLLED TRIALS

Hamid Reza Hosseini, DDS, PGDip

Professor: Athanasios E. Athanasiou

Assistant Professor: Eleftherios G. Kaklamanos

Aim:

To investigate the effectiveness of pre-surgical infant orthopedics in patients with non-syndromic cleft lip and/or palate and evaluate the quality of the available evidence.

Materials and methods:

An electronic search without restrictions for published and unpublished literature, together with hand searching, was carried out. Randomized controlled trials (RCTs) investigating the effects of pre-surgical infant orthopedic appliances reviewed. The risk of bias was assessed using the Cochrane Collaboration's Risk of Bias assessment tool for RCTs and the quality of evidence assessed according the Grades of Recommendation, Assessment, Development and Evaluation (GRADE) approach.

Results:

The initially identified 1,043 records were finally reduced to 20 full-text reports concerning a total of 118 patients with unilateral complete cleft lip and palate and 16 with cleft of the soft and at least two thirds of the hard palate. Eighteen of the eligible records comprised a part of a larger trial. Eight publications were considered as being of low, four of unclear and eight of high risk of bias.

In general, the investigated appliances did not present significant effects when compared to each other or to no treatment in terms of feeding characteristics and general body growth, facial esthetics, cephalometric variables, maxillary dentoalveolar variables and dental arch relationships, speech and language related variables, caregiver-reported outcomes, economic evaluation related outcomes, as well as, adverse effects and problems related to the appliances or the applied procedures. Overall, the quality of the available evidence was considered low.

Conclusions:

The aforementioned findings could provide initial guidance in the clinical setting. However, given the multitude of parameters which may have affected the results, good practice would suggest further research in the respective field, in order to arrive at more robust relevant recommendations for management decisions in individual cases.

DEDICATION

I would like to dedicate this thesis to my parents.

DECLARATION

I declare that all the content of the thesis is my own work. There is no conflict of interest with any other entity or organization.

Name: Hamid Reza Hosseini

Signature:

ACKNOWLEDGEMENTS

I would like to express my special thanks, and indebtedness to my mentor, Dean, Professor and Program Director in Orthodontics Athanasios E. Athanasiou, who continually and convincingly conveyed a spirit of adventure in regards to orthodontics and research, and provided me with vast and invaluable support and guidance during my years as an orthodontic resident.

I would like also to express my deepest appreciation to my supervisor, Assistant Professor in Orthodontics Eleftherios G. Kaklamanos; without his guidance and perseverance this dissertation would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1. INTRODUCTION

Clefts of the upper lip, alveolar ridge and palate are considered to be one of the most common congenital malformations in humans (Mossey and Little, 2002). These defects, involving various soft and osseous tissues of the oral cavity, occur when the morphogenesis of the upper lip and the palate taking place between the 6th and the 12th week of fetal development, deviates from normal and the fusion of the maxillary with the two medial nasal prominences followed by the primary palate with the two lateral palatal shelves is disrupted (Sperber and Sperber, 2013). Such cleft defects may have manifold significant detrimental consequences for the affected individual and their family environment when considering its potential impact on normal feeding and bodily growth, proper hearing and speech development, harmonious growth of the face and occlusion together with the various psychosocial parameters associated with the deformative nature of the condition (Berkowitz, 2013a).

The use of pre-surgical infant orthopedics was introduced as a means to improve management and treatment outcomes of patients with cleft lip and/or palate in the short-term by assisting feeding, preventing delays in development and helping in normalizing the function of deglutition; the resulting amelioration in the distorted dental arch and nasal forms also facilitates surgical procedures by minimizing tension at the surgical site (Gnoinski et al., 1990). In the long-term, the proposed advantages include enhanced speech development as well as improved maxillofacial growth and facial esthetics; with the assumption that these will decrease the need for specialist intervention in the future (Prahl et al., 2001).

During the last forty years, such procedures became integrated into the comprehensive care protocols for patients with clefts of many teams around the world (Shaw et al., 1992; Prahl et al., 2001; Lohmander et al., 2004; Masarei et al., 2007; Long et al., 2011). Interestingly however, this practice still remains a subject of controversy (Kuijpers-Jagtman and Prahl, 2013) and the relevant

data have not been correlated regarding the quality of evidence to provide critical insights into the advisability of the relevant practices/recommendations. When a child is born with a cleft, parents are naturally highly concerned to know whether any interventions are possible that may enhance feeding, speech, appearance or other outcomes. Pre-surgical infant orthopedics may increase the already significant burden of care during a period where caregivers and families are attempting to cope with the stressful situation of having a child with a cleft defect (Solnit and Stark, 1961). Thus, the World Health Organization has suggested evaluating treatment modalities in the field of craniofacial abnormalities should be a continuous activity to alleviate their impact on individuals, families and healthcare systems (World Health Organization, 2002).

The aim of this thesis, presented to the Hamdan Bin Mohammed College of Dental Medicine of the Mohammed Bin Rashid University of Medicine and Health Sciences in partial fulfillment of the requirements for the Degree of Master of Science in Orthodontics, was to systematically investigate the effectiveness of pre-surgical infant orthopedics in patients with non-syndromic cleft lip and/or palate and evaluate the quality of the available evidence.

2. REVIEW OF THE LITERATURE

Clefts of the lip and/or palate constitute congenital defects of the superior part of the oral cavity occurring between the 6th and the 12th week of fetal development (Sperber and Sperber, 2013). Failure of fusion between the two bilateral maxillary processes with the two medial nasal prominences results in disruption of normal morphogenesis of the upper lip and primary palate area; any impairment in the union between the primary palate and two lateral palatal shelves leads to the creation of varying degrees of communication between the oral cavity on the one part and the nasal cavity and the superior pharyngeal area on the other (Sperber and Sperber, 2013).

Based on the embryological, anatomical, and physiological characteristics of the defects, the variations in clefts of the lip and the palate may categorized into four general classifications: clefts involving the lip and the alveolus, clefts involving the lip and the palate, clefts of the palate alone, and finally congenital insufficiency of the palate. In addition, clefts can occur incompletely or completely and present unilateral or bilateral localization (Berkowitz, 2013b).

2.1. Epidemiology

Cleft defects most probably predate modern humans themselves, but early documentation is sparse. During the second century BCE, the Greek physician Galen, made a reference to cleft lip and used the term *κολοβώματα* (colobomata – tissue defects) to describe the malformation (Millard, 1976).

Based on contemporary epidemiological data, clefts constitute one of the most common malformations occurring in humans (Mossey and Little, 2002). Overall, 85% of orofacial clefts are considered to be non-syndromic, with the remaining 15% involving almost 300 different known syndromes (Online Mendelian Inheritance in Man, OMIM, 2000). Based on international data, the overall proportion of clefts has been calculated to be around 1.47 per 1,000 live births

(Panamonta et al., 2015) (Table1).

However, significant variability exists depending on the geographical region concerned and the ethnic origin of the studied population. A higher proportion of clefts per 1,000 live births has been observed in Asia, North America and Europe, and lower rates in Africa (Table 1). Some studies have recorded higher rates existing in Asian and North American native populations, with Caucasian populations having an intermediate prevalence and the lowest prevalence of cleft defects existing in African populations (Gorlin et. al., 2001). However, the relatively low proportion of clefts in those countries with poor birth defects surveillance has raised questions concerning relative ascertainment of these cases (Mossey and Modell, 2012).

Table 1. Geographical variation in cleft lip or cleft lip with cleft palate and isolated cleft palate occurrence per 1000 live births by continent (Panamonta et al., 2015).

Continent	Numbers of live births	Proportion of clefting per 1000 live births (95% CI)
Asia	9,965,084	1.57 (1.54-1.60)
North America	11,728,914	1.56 (1.53-1.59)
Europe	3,236,253	1.55 (1.52-1.58)
Oceania	2,125,912	1.33 (1.30-1.36)
South America	3,229,179	0.99 (0.96-1.02)
Africa	380,273	0.57 (0.54-0.60)
Total	**30,665,615**	**1.47 (1.44-1.50)**

95% CI: 95% Confidence Interval

According to European and US studies, the most frequently occurring type of clefts in general is the unilateral cleft lip and palate defect that accounts for 30-35% of non-syndromic clefts, followed by cleft lip and cleft palate each 20-25%, and bilateral cleft lip and palate with about 10%. Submucous and other types of cleft malformations account for the remainder (Hagberg et al., 1997). In the cleft lip and palate group, 80% are unilateral and 20% bilateral type of defects (Mossey and Modell, 2012).

It is now generally accepted that males have a greater tendency towards cleft lip and/or palate and females toward isolated cleft palate deformities. However, the gender ratio differs depending on the severity of the cleft (Mossey et al., 2002), associated malformations, the number siblings affected, ethnicity and paternal age (Mossey et al., 2009). No unequivocally acceptable explanation for differences between genders exists, but differences in the critical stages of craniofacial development timing are believed to be involved (Burdi et al., 1969).

2.2. Etiology

The genetic and phenotypic heterogeneity observed in non-syndromic cleft defects has made the identification of specific etiological factors difficult. Nowadays, it is broadly recognized that both genetic and environmental parameters play a role in the condition and most cases are considered to be polygenic and/or multifactorial in origin (Cobourne, 2004; Vieira, 2012).

Adringer and et al. (1989) were the first to describe an association between Transforming Growth Factor-α gene (*TGFA*) and non-syndromic cleft lip and palate in a Caucasian population. Variants of the Interferon Regulatory Factor 6 (*IRF6*) gene have also been implicated in the formation of oro-facial clefts (Zucchero et al., 2004). In addition, the Msh Homeobox 1 (*MSX1*) gene, a member of distinct subfamily of homeobox genes related to the Dorosophila msh (muscle segment homebox) gene (Holland, 1991), has been proposed as another strong candidate for involvement in orofacial clefts and dental abnormalities (van den Boogaard et al., 2000), as it is possibly associated with lack of adequate palatal mesenchyme (Satokata et al., 1994). More recent genetic studies have identified a constellation of different gene variants associated with cleft lip and palate defects (Rahimov et al., 2011; Weatherley-White et al., 2011), leading to the realization that the original quest for major gene mutations capable of explaining most cases of cleft defects is unlikely to be successful (Viera, 2012).

The possibility exists that genetic factors such as these may still be under the influence of parameters associated with the environment (Viera, 2012). Many studies have reported an association of cleft defects with environmental factors such as anti-folate drugs, retinoids, hypoxia, increased temperature and Vitamin B6 deficiency (Murray, 2002; Little et al., 2003; De La Vega and Martinez, 2006; Mossey et al., 2009; Dixon et al., 2011). However, the most important environmental influences accepted as associated with cleft lip and palate are maternal cigarette smoking, alcohol consumption and multivitamin use (Viera, 2012). Maternal smoking can lead to embryonic hypoxia and an increased incidence of non-syndromic cleft-lip and palate (Wyszynski et al., 1997; Viera, 2008). Women who abuse alcohol during their pregnancy are at high risk of bearing a child with cleft manifestation, especially if their offspring carries a specific genetic variation in ADH1C (Alcohol Dehydrogenase 1C) (Boyles et al., 2010). On the contrary, maternal multivitamin use has been suggested to decrease the susceptibility to cleft malformations through interaction with *IRF6* (Wu et al., 2010) and *ZNF533* genes (Wu et al., 2011).

2.3. Detrimental effects of clefts of the lip and the palate

Clefts of the lip and the palate may have a multitude of significant detrimental consequences for the affected individual, as well as for their family environment, given that such defects can affect normal feeding and bodily growth, proper hearing and speech development, harmonious facial and occlusion growth together with various psychosocial parameters associated with the deformative nature of the condition (Berkowitz, 2013a).

Feeding constitutes an extremely important aspect concerning parents and potentially increasing the burden of care for the families (Adams et al., 1999). Problems with food intake may adversely affect nutritional status early in life for children with cleft lip and palate or cleft palate leading to lighter and shorter infants than unaffected infants (Avedian 1980; Duncan 1983; Becker 1998). Although some research has indicated that children with clefts attain their expected weight and

height by about 2 years of age (Lee, 1997), other studies have reported delayed skeletal maturity in boys with clefts over a period from 6 to 20 years (Jensen, 1983). Additional developmental delays have been observed in tests regarding cognition, comprehension and expressive language abilities (Jocelyn, 1996), as well as fine motor and gross motor skills (Neiman, 1997).

There is a significant risk of hearing loss in infants born with cleft lip and palate. Dysfunction of the eustachian tube and effusions of middle ear have been observed to be common among these patients (Paradise et al., 1969). Speech sound acquisition and a child`s ability to benefit from speech therapy can be affected by hearing loss parameters adding to the problems experienced by these children regarding speech and language development because of the altered anatomical environment (Konst et al., 2000).

Patients with cleft lip and palate do not exhibit harmonious facial and occlusion growth, including decreased upper facial height, retrusive maxilla, tendency to hyperdivergent growth pattern, disruption of dental development, retroclined incisors, obtuse interincisal and nasolabial angles, posterior crossbites and Class III malocclusion (Dahl, 1970; Athanasiou et al., 1986; Ranta, 1986; Athanasiou et al., 1987a; Athanasiou et al., 1987b; Athanasioue et al., 1988; Athanasiou et al., 1990; McCance et al., 1990; Athanasiou et al., 1991a; Athanasiou et al., 1991b; Semb, 1991; Mazaheri et al., 1993; Nollet et al., 2008; Berkowitz, 2013c). These features may be attributed to the intrinsic developmental factors associated with the cleft itself, or related to functional problems caused by compromised nasal breathing, deviated tongue position, muscles abnormalities and inadequate functioning, or possibly even to iatrogenic factors, such as scarring or the effects of orthodontic therapy (Ross, 1987).

The psychosocial effects of the condition are difficult to define and quantify precisely and may not be directly related to the severity of the problem, but more a result of the impairment of function and esthetics (Turner et al., 1998). The multitude of findings such as depression and low self-esteem, unemployment and behavioral problems, suggest that malformation is an aspect

requiring serious consideration when addressing the problems of orofacial clefts (Murthy 2009). Moreover, the anomaly may also affect the parents psychologically, as they are concerned about the disturbances in function and esthetics (Solnit and Stark, 1961; Riski, 1991).

2.4. Treatment protocols for cleft lip and palate patients

The various complex problems facing an individual with cleft lip and/or palate necessitates an interdisciplinary team care approach to enable the child to best overcome the shortcoming of the disturbances that occur in their growth and development and become a functioning and contributing member of the society. The American Cleft Palate Craniofacial Association, stated that ‘’a team consists of an operating surgeon, orthodontist, speech language pathologist, and at least one additional specialist from otolaryngology, audiology, pediatrics, genetics, social work, psychology, and general pediatric or prosthetic dentistry, who meet face to face to evaluate and produce treatment plans for the patients, is required to attend this complex anomaly’’ (American Cleft Palate Craniofacial Association, 2009).

Specific treatment protocols exhibit significant variation between different therapeutic centers and teams, especially with regards to the parameter of surgery (Shaw et al, 2000). Tables 2 and 3 provide examples of treatment protocols for cleft lip and palate patients.

Table 2. Example of a treatment protocol for cleft lip and palate patients (Mølsted, 1992).

Timing	
At birth	Counselling and information for parents. Neonatal repair undertaken by some. Regular hearing/ENT checks until adulthood.
3 months	Lip repair.
6 months	Palate repair (Von Langebeck).
2-10 years	Pharyngoplasty if needed to improve velopharyngeal competence. Further advice on oral health and regular care. Speech assessment and treatment as required.
7-10 years	Secondary alveolar bone grafting. Preparation for alveolar bone grafting. Expansion (Quad/Tri-helix) and anterior alignment if necessary. Stabilization of arch form. Maintenance of oral hygiene.
11-15 years	Conventional orthodontic treatment.
18+ years	Osteotomy if necessary. Rhinoplasty if necessary. Presurgical orthodontic if necessary. Bridge and denture work as necessary.

Table 3. Example of a treatment protocol for cleft lip and palate patients (Berkowitz, 2013).

Timing	
3 months	Lip adhesion.
10 months	Definitive lip surgery (rotation advancement).
18-24 months	Hard and soft palate closure (von Langenbeck with vomer flap).
5-7 years	Orthodontic expansion (quad helix).
6-8 years	Superior-based pharyngeal flap (if necessary).
7-9 years	Bone graft (iliac crest).
8 or later	Protraction facial mask (if necessary).
Variable	Maxillary surgical advancement either Le Fort I or distraction osteogenesis. Lip/nose revisions techniques.

2.5. Pre-surgical infant orthopedics

The concept of pre-surgical infant orthopedics was introduced as a means to improve management and treatment outcomes of patients with cleft lip and/or palate (Kuijpers-Jagtman and Prahl, 2013). The objectives of early active maxillary orthopedics have been suggested as two-fold. First, in cases of wide unilateral clefts or protrusive pre-maxilla in bilateral clefts, initial surgical lip closure could be difficult because of the distance the tissue must be mobilized to close the defect. This may cause excessive tension on the surgical site, which could lead to wound dehiscence (Berkowitz, 2013a). Pre-surgical infant orthopedics protocols are intended to allow an earlier, more ideal lip closure with minimum tissue tension since the soft tissues overlie a more normal bony anatomy (Sierra et al., 1995). Secondly, pre-surgical infant orthopedics is seen as a means to prevent the collapse of lateral alveolar segments, which would usually occur in the absence of intervention, leading to mal-alignment. With early orthopedic intervention, a more normal arch form could be achieved, resulting in enhanced alignment of the segments and again a more normal or near normal anatomical relationship in the area of the cleft defect (Sierra et al., 1995).

2.5.1. Introduction to pre-surgical infant orthopedics

Before the advent of contemporary pre-surgical infant orthopedics protocols, Brophy (1927) developed the practice of wiring together both ends of the cleft; these wires were later tightened in order to reduce the width to be dealt with by the subsequent lip surgery. Later, adhesive tape strapping was used to decrease the width of the clefts (Winters and Hurwitz, 1995). In the 1950s, McNeil employed Scott`s theory (Scott, 1953) as a justification for his pre-surgical orthopedics protocol. It was suggested that after the downward and forward growth of the nasal septum the segments of the palate would detach and remain deficient and cause the mid-face deficiency that is usually a characteristic of patients with cleft lip and palate. McNeil contended that by having the palatal segments molded in the anatomical correct position, normal growth of the mid-face

could be achieved while reducing the width of the cleft (McNeil 1950; 1954). During the 70s and 80s, Hotz popularized the use of plates without use of active elements with the intention of exploiting the intrinsic developmental processes in the maxillary region during the infantile period (Hotz and Gnoinski 1976; Hotz et al. 1986). A combined protocol of presurgical orthopedics and bone grafting was also advocated during the same period (Pruzansky, 1964).

More recently, the nasoalveolar molding (NAM) technique was described (Grayson et al., 1999*;* Maull et al., 1999, Grayson et al., 2001) to pre-surgically mold the lip, the alveolus and the nasal cartilage in infants born with cleft lip and palate. The concept of this protocol is based on Matsuo's principle. Due to the increased level of plasticity concurrent with the increased levels of maternal estrogen, initiating the procedure within the first 6 weeks of life allows the nasal cartilage to be molded effectively (Matsuo et al., 1989).

2.5.2. Types of pre-surgical infant orthopedic appliances

Pre-surgical infant orthopedic appliances can either be constructed so as to apply a kind of active force system on the alveolar segments to move them into the desired direction or, alternatively, passively encourage cleft segment alignment during maxillary growth (Kuijpers-Jagtman and Prahl, 2013).

In the first category the applied force may result from activated springs, screws and elastic chains (DiBiase and Hunter, 1983; Millard et al., 1999, Moore et al., 2005; Cash, 2012; Berkowitz and Mejia, 2013). If desirable, anchorage for these appliances may be obtained by placing pins into the palate (Georgiade 1964; Georgiade et al., 1968; Latham 1980). Alternatively, active appliances may be fabricated on sequential sectioned and adjusted maxillary casts, where the alveolar segments are gradually repositioned into more normal positions, thus exerting forces on the infantile maxillary alveolar ridges when the active appliances are placed in the oral cavity

(Kuijpers-Jagtman and Prahl, 2013). External strapping across the cleft segments may constitute part of the treatment protocol (McNeil, 1950).

The so-characterized passive pre-surgical infant orthopedic appliances are meant to encourage spontaneous cleft segment alignment during maxillary growth by manipulating either the shape of the alveolar processes in the laboratory cast or by selectively removing and adding acrylic directly on the plate itself (Zarrinnia et al., 1993; Cash, 2012; Kuijpers-Jagtman and Prahl, 2013). In the latter cases the acrylic is generally removed from the areas where growth of the alveolus is attempted and added onto the opposite side. Such plates may be retained only by suction and adhesion (Hotz, 1969; Hotz and Gnoinski 1976). Extraoral strapping may again be part of the applied protocol (Huddart 1967).

The pre-surgical infant orthopedic appliances used in the nasoalveolar molding technique differ from those used in other protocols by the addition to the acrylic plate of nasal stents to modify the distorted nasal area. The intraoral acrylic plate is held in place with extraoral tape and elastics (Grayson et al., 2008). The stents are added after the palatal segments have been approximated and the premaxilla retracted to sufficiently reduce the pre-existing nasal deformity and enable the start of a more accurate nasal molding (Barillas et al., 2009, Santiago and Grayson, 2009). External force can be applied by external taping of the lip, a head cap with elastic straps across the prolabium, or a surgical lip adhesion (Grayson et al., 2008).

2.5.3. Controversy regarding the use of pre-surgical infant orthopedics

Pre-surgical infant orthopedics has come to be part of comprehensive care protocols for patients with clefts for several surgical teams around the world (Shaw et al., 1992; Prahl et al., 2001; Lohmander et al., 2004; Masarei et al., 2007; Long et al., 2011).

Those supporting its use contend that, in the short-term, they assist in feeding and help avoid delays in development, help in normalizing the function of deglutition, improve the shape of the

distorted maxillary alveolar arch and nasal area form and facilitate surgical procedures by minimizing tension at the surgical site (Gnoinski et al., 1990). In addition, long-term advantages are claimed to include enhanced speech development, as well as, maxillofacial growth and facial esthetics which are assumed to reduce the need for future specialist intervention (Ross and McNamera, 1994; Kuijpers-Jagtman and Ross, 2000; Mishima et al., 2000; Prahl et al., 2001). Advocates of pre-surgical nasolalveolar moulding appliances have argued that, beside other advantages of conventional plates, additional benefits include improvement of nasal symmetry and lip esthetics whilst elongating the columella and correcting nasal cartilage deformity (Grayson et al., 1999; Grayson and Cutting, 2001; Liou et al., 2004).

On the other hand, opponents of the pre-surgical infant orthopedic protocols claim that not only are these approaches expensive and complicated, but they can also have an adverse effect on maxillary growth and speech development (Bardach et al., 1984; Witzel et al., 1984; Kramer et al., 1993).

The present situation is that the application of pre-surgical infant orthopedics protocols remains a subject of strenuous debate among specialists (Kuijpers-Jagtman and Prahl, 2013). Two systematic reviews evaluating pre-surgical infant orthopedics in general (Uzel and Alparslan, 2011; Papadopoulos et a., 2012), and on assessing growth and development outcomes after various feeding interventions in infants with cleft lip and/or palate (Bessell et al., 2011) have been published recently. However, neither of these attempted to ascertain and summarize the available quality of evidence and thus provide an insight on the strength of the relevant recommendations.

Nevertheless, when a baby is born with a cleft, parents are inevitably anxious to know whether any kind of intervention can enhance feeding, speech, appearance or other outcomes. Pre-surgical infant orthopedics may increase the already significant burden of care at a time where caregivers and families are attempting to cope with the situation of having a child with a cleft defect (Solnit and Stark, 1961). Thus, the World Health Organization has suggested that evaluation of treatment

modalities in the area of craniofacial abnormalities should be a continuous activity in order to reduce their impact on individuals, families and healthcare systems (World Health Organization, 2002).

3. AIM

3.1. Aim of the systematic review

To investigate the effectiveness of presurgical infant orthopedics in patients with non-syndromic cleft lip and/or palate and evaluate the quality of the available evidence.

3.2. Objectives of the systematic review

a. To examine the effectiveness presurgical infant orthopedics compared to no treatment regarding outcomes relevant to feeding characteristics and nutritional status, facial esthetics, dentofacial cephalometric variables, maxillary dentoalveolar variables, dental arch relationships, hearing, speech and language evaluation, patient and caregiver-reported outcomes, economic evaluation related outcomes, as well as, adverse effects and problems related to PSIO appliances and procedures.

b. To examine the comparative effectiveness of various PSIO appliances and modalities regarding outcomes relevant to feeding characteristics and nutritional status, facial esthetics, dentofacial cephalometric variables, maxillary dentoalveolar variables, dental arch relationships, hearing, speech and language evaluation, patient and caregiver-reported outcomes, economic evaluation related outcomes, as well as, adverse effects and problems related to PSIO appliances and procedures.

3.3. Null hypotheses

a. There is no difference between the outcomes of presurgical infant orthopedics compared to the outcome of no treatment regarding outcomes relevant to feeding characteristics and nutritional status, facial esthetics, dentofacial cephalometric variables, maxillary dentoalveolar variables, dental arch relationships, hearing, speech and language evaluation, patient and

caregiver-reported outcomes, economic evaluation related outcomes, as well as, adverse effects and problems related to PSIO appliances and procedures.

b. There is no difference in the comparative effectiveness of various PSIO appliances and modalities regarding outcomes relevant to feeding characteristics and nutritional status, facial esthetics, dentofacial cephalometric variables, maxillary dentoalveolar variables, dental arch relationships, hearing, speech and language evaluation, patient and caregiver-reported outcomes, economic evaluation related outcomes, as well as, adverse effects and problems related to PSIO appliances and procedures.

4. MATERIALS AND METHODS

4.1 Protocol Development

The present review was based on a specific protocol developed following the guidelines outlined in the PRISMA statement (Moher et al., 2009) and the Cochrane Handbook for Systematic Reviews of Interventions (version 5.1.0) (Higgins and Green, 2011).

The present protocol was registered with PROSPERO - International Prospective Register of Systematic Reviews, produced by the Centre for Reviews and Dissemination (CRD) at the University of York, United Kingdom (UK), and funded by the National Institute for Health Research (NIHR), UK. This protocol is available freely online in the PROSPERO registry website (see Appendix I, Hamid Reza Hosseini, Eleftherios G. Kaklamanos, Athanasios E. Athanasiou. Treatment outcomes of pre-surgical infant orthopedic appliances in patients with non-syndromic cleft lip and/or palate: a systematic review and meta-analysis of randomized controlled trials. PROSPERO 2016: CRD42016047940; Available from http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID= CRD42016047940).

4.2. Selection criteria applied for the review

The selection criteria for the domains of study design, participant characteristics, intervention characteristics and principal outcome measures applied for the present review were as follows:

4.2.1. Types of study design

Studies included in the present thesis had to be Randomized Clinical Trials (RCTs) evaluating outcomes of PSIO appliance treatment. Animal studies, non-comparative studies (case reports and case series), systematic reviews and meta-analyses were excluded.

The type of study design was assessed using the algorithm available from SIGN (Scottish Intercollegiate Guidelines Network) available from http://www.sign.ac.uk (Appendix II).

4.2.2. Types of participants

The included studies could involve children of any age with any kind of non-syndromic cleft lip and/or palate defect.

4.2.3. Types of interventions

The included studies could involve patients using any type of PSIO appliance compared to a non-PSIO treated or an alternative PSIO appliance treated control group.

4.2.4. Types of outcome measures

The studies included in the present review had to provide quantified measurements on outcomes relevant to PSIO appliance treatment.

4.3. Search strategy for identification of studies

The principal investigator (HRH) developed detailed search strategies for each database searched. These were based on the strategy developed for MEDLINE but revised appropriately for each database to take account of the differences in controlled vocabulary and syntax rules. The following electronic databases were searched (Appendix III): MEDLINE via PubMed (http://www.ncbi.nlm.nih.gov/pubmed), the Cochrane Central Register of Controlled Trials (CENTRAL) (http://onlinelibrary.wiley.com/cochranelibrary), Scopus (www.scopus.com), Web of Science™ Core Collection (http://apps.webofknowledge.com/), Latin-American and Caribbean System on Health Sciences Information (LILACS) (http://lilacs.bvsalud.org/en/), National Databases of Indian Medical Journals (IndMed) (http://indmed.nic.in/indmed.html) and Arab

World Research Source (http://0-web.a.ebscohost.com.amclb.iii.com). Unpublished literature was accessed electronically using Google Scholar (https://scholar.google.com), ClinicalTrials.gov (http://clinicaltrials.gov), International Standard Randomised Controlled Trial Number (ISRCTN) registry (http://www.isrctn.com) and OpenGrey (http://www.opengrey.eu). In addition, Pro-Quest Dissertation and Theses Global database (http://search.proquest.com) was searched.

No restriction was placed on the language, date or status of publication. In addition, efforts were made to obtain conference proceedings and abstracts where possible and the reference lists of all eligible studies for additional records were searched. Where additional information for some publication was to be needed, we would try to contact the respective authors.

4.4. Selection of studies and data extraction

The principal investigator (HRH) and the thesis supervisor (EGK) assessed the retrieved records for inclusion independently. They were not blinded to the identity of the authors, their institution, or the results of the research. They obtained and assessed, again independently, the full report of records considered by either reviewer to meet the inclusion criteria. Disagreements were resolved by discussion or consultation with the thesis co-supervisor (AEA). A record of all decisions on study identification was kept.

The same two persons performed data extraction independently and any disagreements were again resolved by discussion or consultation with the thesis co-supervisor (AEA). Data collection forms were used to record the desired information.

a. Bibliographic details of the study.

b. Details on study design and verification of study eligibility.

c. Participant characteristics (where available number, age, gender).

d. Intervention characteristics (PSIO appliance used, PSIO treatment protocol).

e. Details on outcomes assessed and assessment procedures.

f. Additional information: a prior sample size calculation, methodology reliability assessment and data on compliance with the assigned intervention protocol.

If clarifications were needed on the published data or additional material then attempts to contact the corresponding authors would be made.

The outcomes relevant to PSIO appliance treatment retrieved from the studies included in the present review were categorized as follows:

a. Feeding characteristics and nutritional status.

b. Facial esthetics.

c. Dentofacial cephalometric variables.

d. Maxillary dentoalveolar variables.

e. Dental arch relationships.

f. Hearing, speech and language evaluation.

g. Patient and caregiver-reported outcomes.

h. Economic evaluation related outcomes.

i. Adverse effects and problems related to PSIO appliances and procedures.

4.5. Data synthesis and assessment of publication bias

In situations where the retrieved data used different indices measuring the same concept on different scales with a high degree of correlation, the effects of the interventions were planned to be expressed as standardized values (i.e. the Standardized Mean Difference (SMD) together with the relevant 95% Confidence Interval (CI)), in order to enable quantitative synthesis (Deeks et al., 2001). In case that in a particular comparison only one index was recorded, the intervention effect was planned to be expressed as the Weighted Mean Difference (WMD) together with the 95% CI. The random effects method for meta-analysis was to be used to combine data from studies that reported similar measurements in appropriate statistical forms (Der Simonian and Laird, 1986,

Borenstein et al., 2009), since they were expected to differ across studies due to clinical diversity, in terms of participant and intervention characteristics.

To identify the presence and extent of between-study heterogeneity, the overlap of the 95% CI for the results of individual studies was to be inspected graphically, and Cochrane's test for homogeneity and the I^2 statistic were to be calculated (Higgins and Green, 2011). The results of the I^2 statistic were to be interpreted as follows (Higgins and Greene, 2011):

- I^2 from 0% to 40%: heterogeneity might not be important;
- I^2 from 30% to 60%: may represent moderate heterogeneity;
- I^2 from 50% to 90%: may represent substantial heterogeneity;
- I^2 from 75% to 100%: considerable heterogeneity.

If deemed possible, exploratory subgroup analyses were planned according to participant and intervention characteristics. In addition, if a sufficient number of trials were identified, analyses were planned for "small-study effects" and publication bias (Higgins and Green, 2011).

All analyses were to be carried out with Comprehensive Meta-analysis software 2.2.046 (©2007 Biostat Inc.). Significance (a) was set at 0.05, except for 0.10 used for the heterogeneity tests (Ioannidis, 2008).

4.6. Risk of bias assessment and determination of the quality of evidence

The principal investigator (HRH) and the thesis supervisor (EGK) were to assess the risk of bias in the included studies, independently and in duplicate during the data extraction process, using The Cochrane Collaboration's Risk of Bias assessment tool for RCTs (Higgins and Green, 2011). Any disagreements were to be resolved by discussion or consultation with the thesis co-supervisor (AEA). The Risk of Bias assessment tool includes the following domains.

a. Random sequence generation (selection bias).

b. Allocation concealment (selection bias).

c. Blinding of participants and personnel (performance bias).

d. Blinding of outcome assessors (detection bias).

e. Incomplete outcome data (attrition bias).

f. Selective outcome reporting (reporting bias).

g. Other sources of bias.

After entering in the data extraction form the information reported in each study, every domain would receive a judgment of low, high or unclear risk of bias (indicating either lack of sufficient information to make a judgment or uncertainty over the risk of bias) (Higgins and Green, 2011). Subsequently, studies were to be judged as being of low, unclear or high risk of bias.

a. Low risk of bias (plausible bias unlikely to seriously alter the results): all key domains of the study are at low risk of bias.

b. Unclear risk of bias (bias that raises some doubt about the results): one or more key domains of the study are unclear.

c. High risk of bias (bias that seriously weakens confidence in the results): one or more key domains of the study are at high risk of bias.

The quality of evidence and strength of recommendations at longest follow up available for key outcomes of the systematic review were ultimately to be assessed based on the Grades of Recommendation, Assessment, Development and Evaluation (GRADE) approach (Guyatt et al., 2011). The GRADE profiler (GRADEpro) software (available www.gradepro.org; © 2015, McMaster University and Evidence Prime Inc.) was to be used to facilitate the summary regarding the quality of evidence using the GRADE approach. The principal investigator (HRH) and the thesis supervisor (EGK) were to assess the quality of evidence independently and in duplicate. Any disagreements were to be resolved by discussion or consultation with the thesis co-supervisor (AEA).

For the purpose of summarizing risk of bias across studies, where possible, relevant information was to be judged as being of low, unclear or high risk of bias.

a. Low risk of bias: most information is from studies at low risk of bias.

b. Unclear risk of bias: most information is from studies at low or unclear risk of bias.

c. High risk of bias: information from studies at high risk of bias could have an effect on the interpretation of the results.

5. RESULTS

5.1. Results of the search

The flow of records through the reviewing process is shown in Figure 1. Initially 1043 records were identified, and 167 were identified as duplicates and 813 more on the basis of their title and abstract. Finally, 20 full-text reports were included in the systematic review (Severens et al., 1998; Konst et al., 1999; Konst, 2000; Prahl et al., 2001; Konst et al., 2002; 2003a; 2003b; 2003c; Prahl et al., 2003; Bongaarts et al., 2004; Konst et al., 2004; Prahl et al., 2005; Bongaarts et al., 2006; Prahl et al., 2006; Masarei et al., 2007; Bongaarts et al., 2008; Prahl et al., 2008; Bongaarts et al., 2009; Chang et al., 2014; Noverraz et al., 2015). The Konst (2002) publication involves a thesis containing the data available in subsequent publications Konst et al., (1999; 2000; 2003a; 2003b; 2003c; 2004). In the context of the present review it was considered only for the outcome "future need of speech therapy" that was not reported in the abovementioned publications.

5.2. Study characteristics

The characteristics of the studies included in the present systematic review are presented in Tables 4-7. The papers, which were published between 1998 and 2015, recruited 118 patients with unilateral complete cleft lip and palate (UCLP) and 16 with cleft of the soft and at least two thirds of the hard palate. No studies with children with bilateral cleft lip and palate were retrieved.

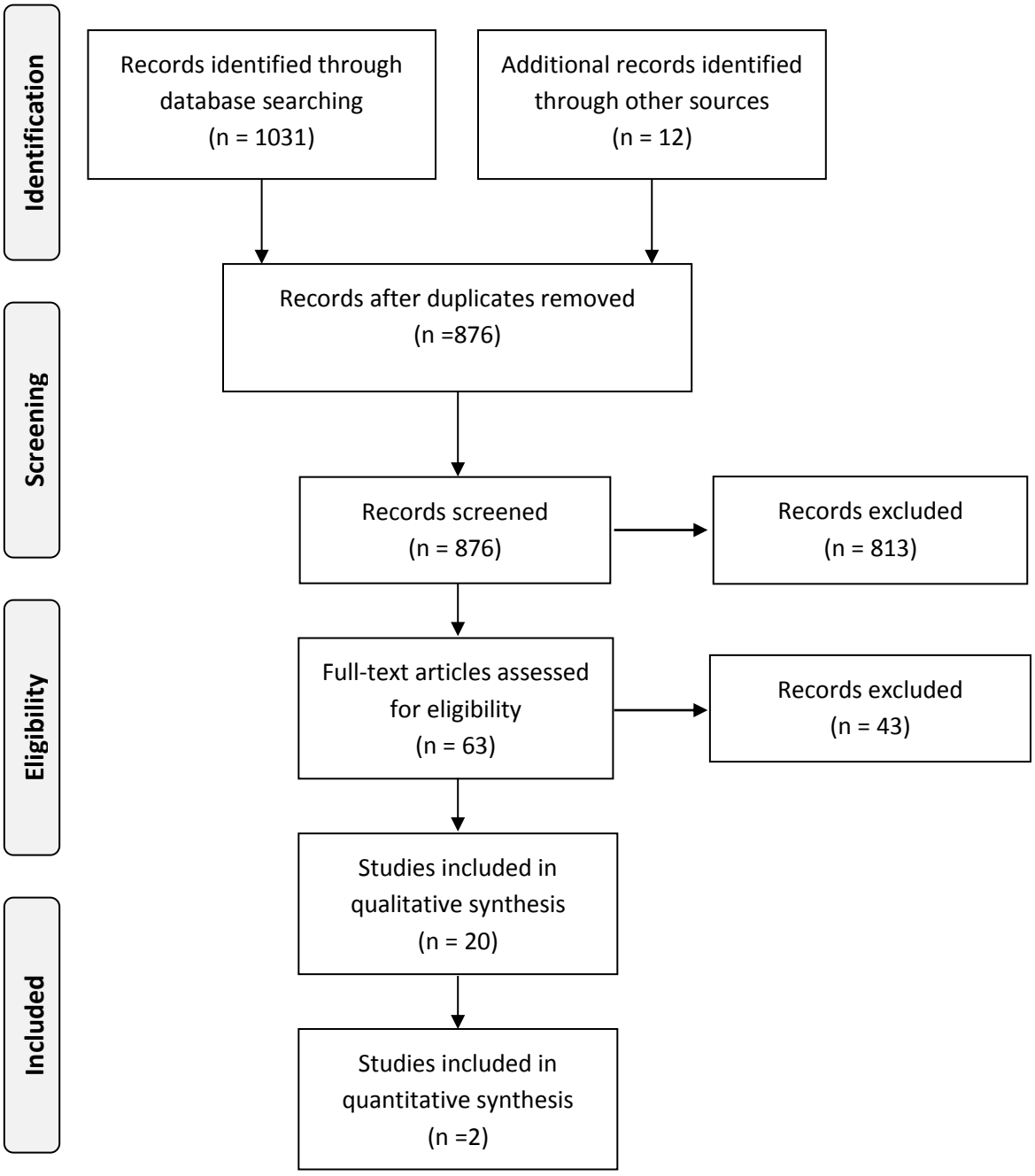


Figure 1. Flow of records through the reviewing process.

Eighteen of the eligible publications comprised a part of a larger trial, the DUTCHCLEFT, that was conducted in three academic cleft palate centers in the Netherlands, enrolled in total 54 patients (41 boys and 13 girls) with non-syndromic UCLP from 1993 to 1996 and compared the effect of a passive PSIO appliance without external retention to no treatment (Severens et al., 1998;

Konst et al., 1999; Konst, 2000; Prahl et al., 2001; Konst et al., 2002; 2003a; 2003b; 2003c; Prahl et al., 2003; Bongaarts et al., 2004; Konst et al., 2004; Prahl et al., 2005; Bongaarts et al., 2006; Prahl et al., 2006; Bongaarts et al., 2008; Prahl et al., 2008; Bongaarts et al., 2009; Noverraz et al., 2015). Subjects randomized into the intervention group, received treatment commencing within 2 weeks post partum till surgical closure of the soft palate (around 52 weeks). The appliance was worn 24 h per day, except for cleaning, and was adjusted, repaired or replaced as needed. Check-up appointments were scheduled every 3 weeks until surgical lip closure (around 18 weeks) and every 4 to 6 weeks thereafter until surgical closure of the soft palate. Subjects belonging to the no-treatment control group did not wear any kind of appliance and were scheduled for appointments at age 6 weeks after birth, as well as, before and after surgical closure of the lip and the soft palate defects.

Two of the 27 children of the intervention group hardly used the appliances, and in one case the appliance was worn for 78 weeks by mistake. These infants remained in their assigned group according to the intention to treat rationale. No other data on whether the plates were worn by every patient as indicated were given.

In the context of the DUTCHCLEFT, in various subsamples from the initial patient group and at various time points until the age of 12 years, the following domains were evaluated:

a. Feeding characteristics and nutritional status (Prahl et al., 2005).

b. Facial esthetics (Prahl et al., 2006; Bongaarts et al., 2008).

c. Dentofacial cephalometric variables (Bongaarts et al., 2009).

d. Maxillary dentoalveolar variables (Prahl et al., 2001; 2003; Bongaarts et al., 2006).

e. Dental arch relationships (Bongaarts et al., 2004; Noverraz et al., 2015)

f. Hearing, speech and language evaluation (Konst et al., 1999; Konst, 2000; Konst et al., 2002; 2003a; 2003b; 2003c).

g. Patient and caregiver-reported outcomes (Prahl et al., 2008)

h. Economic evaluation related outcomes (Severens et al., 1998; Prahl et al., 2001; Konst et al., 2004).

The Masarei et al. (2007) study, conducted in the UK and following a slightly different surgical protocol, involved 33 infants (21 boys and 12 girls) with non-syndromic UCLP and 16 infants (9 boys and 7 girls) with cleft of the soft palate accompanied by at least two thirds of the hard palate. The study compared an active PSIO appliance without external retention to no treatment in the former group and a passive PSIO appliance without external retention in the latter. Subjects randomized into the intervention group, received treatment commencing within 2 weeks post partum till surgical closure of the soft palate (around 6 months). The appliance was worn 24 h per day, except for cleaning, and was adjusted or replaced as needed. Unilateral complete cleft lip and palate subjects were scheduled for check-up appointments 5 or 6 times within the first 3 months (surgical closure of the lip and the anterior palate performed as closely as possible to 12 weeks of age) and another 3 before surgical closure of the soft palate. Patients with cleft palate were seen four times during the first 6 months.

Initially, most of infants (23 out of 25) in the intervention group complied fully with the indicated protocol of appliance wear. However, compliance decreased over time, and only 14 infants were fully compliant during the intervention period and another subject wore the appliance for 12 hours a day.

The following domains were evaluated, at 3 months of age (prior to any surgical intervention) and 12 months of age (approximately 6 months after soft palate surgical repair):

a. Feeding characteristics and nutritional status.

b. Economic evaluation related outcomes.

c. Adverse effects and problems related to PSIO appliances and procedures.

Finally, the Chang et al. (2014) study from Taiwan enrolled 30 infants (19 boys and 11 girls) with complete unilateral cleft lip and palate from May 2013 to March 2013 and compared the effect of the modified Figueroa to the modified Grayson nasolaveolar molding technique. The Figueroa technique uses a PSIO appliance equipped with a nasal stent, without external retention, plus labial tapes. In the Grayson technique the appliance is held in place with the aid of attachment of elastic and surgical tapes. Subjects received treatment commencing within 2 weeks after birth till surgical closure of the lip (around 3 months). The appliance was worn 24 h per day, except for cleaning, and was adjusted every 1 to 2 weeks as needed. All patients completed the trial and compliance was characterized as being exceptionally high. The following domains were evaluated during the initial visit, after nasoalveolar molding but before lip closure surgery, 1 week after surgery and 6 months after surgery:

a. Facial esthetics.

b. Economic evaluation related outcomes

c. Adverse effects and problems related to PSIO appliances and procedures.

Table 4. General characteristics of the studies included in the systematic review – Publications from the DUTCHCLEFT.

Study	Intervention characteristics	Included outcomes	Additional information
DUTCHCELFT Netherlands **[Common items]**	**Group 1**: Passive PSIO appliance without external retention [Zurich / Hotz plate]. **Group 2**: No intervention.	**See details below in individual papers**	**A priori sample calculation:** SNA difference of 3 degrees between the two groups (23 patients per group) **Information on compliance:** Yes
Bongaarts et al., 2004		**5-year-old index** (categorizing arch relationships using reference models)**, Huddart score** (dental arch relationships at the transverse planes)**, overjet, overbite, sagittal occlusion assessment** (scored for deciduous canines and second deciduous molars according to Angle classification) **Assessment:** Impressions taken at 4 and 6 years of age and fabrication of casts	**Reliability of measurements:** Examined
Bongaarts et al., 2006		**Maxillary arch dimensions - arch width, arch depth, arch length, arch form, and vertical position of the lesser segment, variables regarding cleft width** (alveolar cleft width, midpalatal cleft width, posterior cleft width at the tuberosity level), linear arch dimensions (anterior arch width, tuberosity width, total arch depth, total arch length, alveolar cleft margins length) and angular measurements (angulation of the greater alveolar segment in the transverse plane of space, angulation of the smaller alveolar segment in the transverse plane of space, vertical slope of the greater alveolar segment, vertical slope of the smaller alveolar segment). **Contact and collapse of the alveolar segments** in maxillary casts at various ages**.** **Assessment:** Impressions taken at 4, 6 years of age and plaster castes fabricated. (Maxillary casts analyzed using Reflex Microscope, Contact was scored as absent (0) or present (1), collapse scored as absent (0), slight (1), moderate (2), severe (3))	**Reliability of measurements:** Examined. The measurements in the vertical direction had low reliability.
Bongaarts et al., 2008		**Visual analog scores** of facial appearance from full face and cropped photographs (focusing on the mouth and nose area) **Assessment:** 4 and 6 years of age.	**Reliability of measurements:** Examined

Table 4. General characteristics of the studies included in the systematic review – Publications from the DUTCHCLEFT [continued].

Study	Intervention characteristics	Included outcomes	Additional information
Bongaarts et al., 2009		**Dentofacial cephalometric variables evaluation** (angular, linear and ratio variable representing soft and hard tissues, as well as, dental structures) **Assessment:** 4 and 6 years of age. (Lateral head films taken, landmarks digitized on scanned images using Viewbox, version 3.1.0.5, occlusion scored using 5-year old index, esthetics scored using visual analog scale on facial photographs)	**Reliability of measurements:** Examined. The largest errors were found in measurements involving point A or ANS, or the soft tissues. The reliability was good to acceptable, except for two measurements: upper incisor to ANS-PNS angle and ANS-PNS – SN angle. These two measurements were excluded from further analysis.
Konst et al., 1999		**Analysis of prelexical utterances** by means of a perceptually based sensori-motoric classification system. **Assessment:** 12 and 18 months. (Sound production recorded for speech and language evaluation)	**Reliability of measurements:** Examined
Konst, 2000		**Future need for speech therapy** **Assessment:** 2.5 years. (Five trained female speech therapists assessed children`s speech)	**Reliability of measurements:** Examined
Konst et al., 2003a		**Various speech characteristics:** Place of articulation, voice characteristics, nasalization, correctness of articulation, intelligibility, total impression **Assessment:** 2.5 years. (Five trained female speech therapists assessed children`s speech in a blinded perceptual rating procedure on a EAI scales (a seven-point scale) and a 10 point scale for total impression)	**Reliability of measurements:** Examined
Konst et al., 2003b		**Receptive language skills** (Mean length of utterances, mean length of longest utterances) **Assessment:** 2, 2.5, 3 and 6 years. (Reynell Developmental Language Scales Dutch Version test, and in 6-year-olds standardized Dutch language tests)	**Reliability of measurements:** NR
Konst et al., 2003c		**Phonological skills** (Number of acquired consonants, order of phonological development, use of phonological processes, and occurrence of nasal escape) **Assessment: 2, 2.5, 3 years of age.** (A system for assessing phonological development of Dutch children; Fonologische Analyse van het Nederlands: FAN)	**Reliability of measurements:** Examined

Table 4. General characteristics of the studies included in the systematic review – Publications from the DUTCHCLEFT [continued].

Study	Intervention characteristics	Included outcomes	Additional information
Konst et al., 2004		**Cost-effectiveness of PSIO regarding speech intelligibility** **Assessment: 2.5 years of age.** (Intelligibility assessed by five trained speech therapists judging the total impression of speech quality on a 10-point equal-appearing interval scale and costs measured from a social view point in Euro)	**Reliability of measurements:** NR
Noverraz et al., 2015		**Modified Huddart/Bodenham score** **Assessment:** 9 and 12 years of age. (Impressions taken at 9 and 12 years of age and maxillary casts fabricated)	**Reliability of measurements:** Examined
Prahl et al., 2001		**Evaluation of maxillary arch dimensions** (Linear and angular maxillary dentoalveolar variables) **Assessment:** 2, 15, 24, 48, 78 weeks (Impressions taken at 2, 15, 24, 48, 78 weeks and plaster castes fabricated. 13 reference points were described and analyzed using Reflex Microscope technique and 15 dimensions were calculated as well as their increments between the ages)	**Reliability of measurements:** Examined
Prahl et al., 2003		**Presence of contact and/or overlap (collapse) between cleft segments** in maxillary casts at various ages. **Assessment:** 2, 15, 24, 48 and 78 weeks. (Contact was scored as absent (0) or present (1), collapse scored as absent (0), slight (1), moderate (2), severe (3))	**Reliability of measurements:** Examined
Prahl et al., 2005	Feeding instructions given by the orthodontist (use of a squeeze bottle)	**Weight for age, length for age, weight for length, various feeding variables** (time per feeding (min), amount per feeding (mL) and feeding velocity (mL/min)) **Assessment:** 2, 3, 6, 15, 24 weeks (Feeding log/questionnaire were given to the mothers at five different times, weight and length were measured according to national protocol by national infant consultation centers)	**Reliability of measurements:** NR
Prahl et al., 2006		**Visual analog scores** and comparison scores to a reference photograph of facial appearance from full face and cropped photographs (focusing on the mouth and nose area) **Assessment:** 18 months.	**Reliability of measurements:** Examined

Table 4. General characteristics of the studies included in the systematic review – Publications from the DUTCHCLEFT [continued].

Study	Intervention characteristics	Included outcomes	Additional information
Prahl et al., 2008		**Various parameters investigating satisfaction in motherhood** (interaction and caretaking of the baby, coming and goings of the baby, motherhood and life outside and support). **Assessment: 6, 24, 58 weeks.** (Questionnaires, a 4-point scale was used: 1 = very satisfactory to 4 = very unsatisfactory)	**Reliability of measurements:** Examined but the psychometric properties of the instrument not appropriately investigated.
Severens et al., 1998		**Short-term cost-effectiveness of PSIO** (time taken for the surgical lip closure procedure compared to medical and non medical costs until surgical lip closure) **Assessment: 18 weeks.**	***A priori* sample calculation:** NR

NR: Not Reported, EAI: equal-appearing interval

Table 5. General characteristics of the studies included in the systematic review – Remaining studies.

Study	Intervention characteristics	Included outcomes	Additional information
Chang et al., 2004 Taiwan	**Group 1:** Modified Figueroa NAM technique. **Group 2:** Modified Grayson NAM technique.	**Number of clinical visits, total costs, complications, and nasal symmetry** (Nostril height, nostril width, nostril sill height, nostril area) **Assessment:** Initial visit, after NAM but before surgery, 1 week after surgery, 6 months after surgery. (Using standardized craniofacial photographs according to European Association of Cranio-Maxillo-Facial Surgery)	**A priori sample calculation:** Nostril height was considered 1 year after surgery to be the primary outcome and 1-mm difference to be clinically significant. **Reliability measurements:** Examined
Masarei *et al., 2007* UK	**Group 1:** Active or passive plate **Group 2:** No PSIO	**Oral motor skills during feeding** (at 3 months: Neonatal Oral Motor Assessment Scale; at 6 months: Schedule of Oral Motor Assessment)**, physiological measures of bottle feeding/sucking** (at e months Great Ormond Street Measurement of Infant Feeding - length of sucking bursts, peak-to-peak intervals, rate of sucking, suck-swallow rations and percentage pressure generation above baseline pressure in the feeding bottle)**, assessment of pharyngeal stage of swallowing** (at 3 months by videofluoroscopy)**, anthropometry** (at 3 and 6 months: weight, length, head circumference and Body Mass Index) **Assessment:** 3 months, 12 months of age.	***A priori* sample calculation:** Change of 0.8 SD in the anthropometry z scores **Reliability of measurements:** NR

UK: United Kingdom, NR: Not Reported, EAI: equal-appearing interval, NAM: Nasoalveolar molding.

Table 6. Participant characteristics of the studies included in the systematic review – Publications from the DUTCHCLEFT.

Study	Inclusion & exclusion criteria	Number of patients randomized and analyzed
DUTCHCELFT **[General assessment]**	**Inclusion criteria:** complete UCLP, infants born at term, both parents Caucasian and fluent in Dutch language, and trial entrance within 2 weeks after birth. **Exclusion criteria:** other congenital malformations (except for syndactyly) and soft tissue bends.	**Randomized:** 54 children (13 M, 41 F) **Group 1**: 27 (20 M 7 F) **Group 2**: 27 (21 M 6 F)
Bongaarts et al., 2004		**Analyzed at age 4**: **Group 1**: 22 analyzed **Group 2**: 22 analyzed **Analyzed at age 6**: **Group 1**: 22 analyzed **Group 2**: 23 analyzed
Bongaarts et al., 2006		**Analyzed at age 4**: **Group 1**: 23 analyzed **Group 2**: 22 analyzed **Analyzed at age 6**: **Group 1**: 22 analyzed **Group 2**: 23 analyzed
Bongaarts et al., 2008		**Analyzed at age 4**: **Group 1**: 24 analyzed **Group 2**: 21 analyzed **Analyzed at age 6**: **Group 1**: 22 analyzed **Group 2**: 24 analyzed
Bongaarts et al., 2009		**Analyzed at age 4**: **Group 1**: 21 analyzed **Group 2**: 20 analyzed **Analyzed at age 6**: **Group 1**: 21 analyzed **Group 2**: 22 analyzed

Table 6. Participant characteristics of the studies included in the systematic review – Publications from the DUTCHCLEFT.

Study	Inclusion & exclusion criteria	Number of patients randomized and analyzed
Konst et al., 1999		**Analyzed at age 12 months**: **Group 1**: 18 analyzed **Group 2**: 18 analyzed **Analyzed at age 18 months**: **Group 1**: 19 analyzed **Group 2**: 19 analyzed
Konst et al., 2000		**Group 1**: 10 analyzed (8 M, 2 F) **Group 2**: 10 analyzed (9 M, 1 F)
Konst et al., 2003a		**Group 1**: 10 analyzed (8 M, 2 F) **Group 2**: 10 analyzed (9 M, 1 F)
Konst et al., 2003b		**Analyzed at age 2, 2.5, 3**: **Group 1**: 6 analyzed **Group 2**: 6 analyzed **Analyzed at age 6**: **Group 1**: 6 analyzed **Group 2**: 5 analyzed
Konst et al., 2003c		**Analyzed at age 2**: **Group 1**: 9 analyzed **Group 2**: 7 analyzed **Analyzed at age 2.6**: **Group 1**: 9 analyzed **Group 2**: 9 analyzed
Konst et al., 2004		**Analyzed at age 2.5**: **Group 1**: 10 analyzed (8 M, 2 F) **Group 2**: 10 analyzed (9 M, 1 F)
Noverraz et al., 2015		**Analyzed at age 9**: **Group 1**: 24 analyzed **Group 2**: 21 analyzed **Analyzed at age 12**: **Group 1**: 22 analyzed **Group 2**: 22 analyzed

Table 6. Participant characteristics of the studies included in the systematic review – Publications from the DUTCHCLEFT.

Study	Inclusion & exclusion criteria	Number of patients randomized and analyzed
Prahl et al., 2001		**Group 1**: 27 randomized (6 M, 21 F) (24 analyzed) **Group 2**: 27 randomized (7 M, 20 F) (25 analyzed)
Prahl et al., 2003		**At 2 weeks:** **Group 1:** 24 analyzed **Group 2:** 24 analyzed **At 24 weeks:** **Group 1:** 23 analyzed **Group 2:** 24 analyzed **At 58 weeks:** **Group 1:** 16 analyzed **Group 2:** 22 analyzed **At 15 weeks:** **Group 1:** 22 analyzed **Group 2:** 25 analyzed **At 48 weeks:** **Group 1:** 22 analyzed **Group 2:** 24 analyzed **At 78 weeks:** **Group 1:** 19 analyzed **Group 2:** 20 analyzed
Prahl et al., 2005		**At 2 weeks:** **Group 1:** 21 analyzed **Group 2:** 22 analyzed **At 6 weeks:** **Group 1:** 20 analyzed **Group 2:** 24 analyzed **At 24 weeks:** **Group 1:** 20 analyzed **Group 2:** 21 analyzed **At 3 weeks:** **Group 1:** 16 analyzed **Group 2:** 21 analyzed **At 15 weeks:** **Group 1:** 17 analyzed **Group 2:** 24 analyzed
Prahl et al., 2006		**Group 1**: 20 analyzed (17 M, 3 F) **Group 2**: 21 analyzed (17 M, 4 F)

Table 6. Participant characteristics of the studies included in the systematic review – Publications from the DUTCHCLEFT.

Study	Inclusion & exclusion criteria	Number of patients randomized and analyzed
Prahl et al., 2008	**Inclusion criteria:** infants with UCLP or with ICP where the soft palate and at least two thirds of the hard palate was involved. **Exclusion criteria:** infants who required cardiac surgery, neurological impairment, and syndrome known to adversely affect feeding and/ or growth.	**Analyzed at 6 weeks**: **Group 1**: responses from 23 caregivers analyzed **Group 2**: responses from 26 caregivers analyzed **Analyzed at 24 weeks**: **Group 1**: responses from 23 caregivers analyzed **Group 2**: responses from 24 caregivers analyzed **Analyzed at 58 weeks**: **Group 1**: responses from 18 caregivers analyzed **Group 2**: responses from 19 caregivers analyzed [The total number of mothers who had been given the questionnaires were 49]
Severens et al., 1998		**Medical costs** **Group 1**: 23 analyzed **Group 2**: 20 analyzed **Non-medical costs - Travel costs** **Group 1**: 23 analyzed **Group 2**: 20 analyzed **Non-medical costs - Indirect costs** **Group 1**: 15 analyzed **Group 2**: 14 analyzed

M: males, F: females, UCLP: unilateral cleft lip and palate.

Table 7. Participant characteristics of the studies included in the systematic review – Remaining studies.

Study	Inclusion & exclusion criteria	Number of patients randomized and analyzed	
Chang et al., 2014	**Inclusion criteria:** Infants with CUCLP, institutional review board-approved parent- or guardian-signed informed assent. **Exclusion criteria:** Presence of other craniofacial anomalies, incomplete unilateral cleft lip and palate.	**Group 1:** 15 randomized (15 analyzed, 11 M 4 F) **Group 2:** 15 randomized (15 analyzed, 8 M 7 F)	
Massarei *et al.*, 2007	**Inclusion criteria:** infants with UCLP or with ICP where the soft palate and at least two thirds of the hard palate was involved. **Exclusion criteria:** infants who required cardiac surgery, neurological impairment, and syndrome known to adversely affect feeding and/ or growth.	**UCLP:** 33 subjects randomized **Analyzed at 3 months of age:** **Group 1**: 16 analyzed **Group 2**: 16 analyzed **Analyzed at 12 months of age:** **Group 1:** 13 analyzed **Group 2:** 7 analyzed	**ICP:** 16 subjects randomized **Group 1:** 8 analyzed **Group 2:** 8 analyzed **Group 1:** 7 analyzed **Group 2:** 4 analyzed

M: males, F: females, UCLP: Complete unilateral cleft lip and palate.

5.3. Results of risk of bias assessment

Table 8 presents the summary findings of the risk of bias assessment for the included studies. More details can be found in Appendices IV and V. Eight studies were considered as being of low risk of bias (Bongaarts et al., 2004; 2006; 2008; 2009; Prahl et al., 2001; 2003; 2006; Chang et al., 2014), four of unclear risk (Prahl et al., 2005; Masarei et al., 2007; Prahl et al., 2008; Noverraz et al., 2015) and eight of high risk (Severens et al., 1998; Konst et al., 1999; Konst, 2000; Konst et al 2002; 2003a; 2003b; 2003c; 2004).

In general, all studies included in the present review were considered to present low risk of bias regarding the domains of random sequence generation and allocation concealment. Blinding of the participants, caregivers and the personnel providing the instructions was not feasible. However, in the context of the present research design, there was no reason to suggest that bias could be introduced because of absence of blinding in these cases. Moreover, the review authors did not think that bias could be introduced by the methods described in the publications included in the present review regarding blinding of outcome assessment. The examination of the rest of the domains considered produced varying results that contributed to the variable overall assessment of the included studies. More detailed results, according to outcome category, are presented below.

Table 8. Summary of risk of bias assessment.

	Study				
Domain	**Bongaarts et al., 2004**	**Bongaarts et al., 2006**	**Bongaarts et al., 2008**	**Bongaarts et al., 2009**	**Chang et al., 2014**
Random sequence generation	Low	Low	Low	Low	Low
Allocation concealment	Low	Low	Low	Low	Low
Blinding of participants and personnel	Low	Low	Low	Low	Low
Blinding of outcome assessment	Low	Low	Low	Low	Low
Incomplete outcome data	Unclear	Unclear	Unclear	Unclear	Low
Selective outcome reporting	Low	Low	Low	Low	Low
Other potential threats to validity	Unclear	Unclear	Unclear	Unclear	Low
Summary assessment	**Low**	**Low**	**Low**	**Low**	**Low**

Table 8. Summary of risk of bias assessment. [Continued]

	Study				
Domain	**Konst et al., 1999**	**Konst, 2000**	**Konst et al., 2002**	**Konst et al., 2003a**	**Konst et al., 2003b**
Random sequence generation	Low	Low	Low	Low	Low
Allocation concealment	Low	Low	Low	Low	Low
Blinding of participants and personnel	Low	Low	Low	Low	Low
Blinding of outcome assessment	Low	Low	Low	Low	Low
Incomplete outcome data	High	High	High	High	High
Selective outcome reporting	Low	Low	Low	Unclear	Low
Other potential threats to validity	Unclear	Unclear	Unclear	Unclear	Unclear
Summary assessment	**High**	**High**	**High**	**High**	**High**

Table 8. Summary of risk of bias assessment. [Continued]

	Study				
Domain	**Konst et al., 2003c**	**Konst et al., 2004**	**Masarei et al., 2007**	**Noverraz et al., 2015**	**Prahl et al., 2001**
Random sequence generation	Low	Low	Low	Low	Low
Allocation concealment	Low	Low	Low	Low	Low
Blinding of participants and personnel	Low	Low	Low	Low	Low
Blinding of outcome assessment	Low	Low	Low	Low	Low
Incomplete outcome data	High	High	Unclear	Unclear	Low
Selective outcome reporting	Low	Low	Low	Unclear	Low
Other potential threats to validity	Unclear	Unclear	Unclear	Unclear	Unclear
Summary assessment	**High**	**High**	**Unclear**	**Unclear**	**Low**

Table 8. Summary of risk of bias assessment. [Continued]

	Study				
Domain	**Prahl et al., 2003**	**Prahl et al., 2005**	**Prahl et al., 2006**	**Prahl et al., 2008**	**Severens et al., 1998**
Random sequence generation	Low	Low	Low	Low	Low
Allocation concealment	Low	Low	Low	Low	Low
Blinding of participants and personnel	Low	Low	Low	Low	Low
Blinding of outcome assessment	Low	Low	Low	Low	Low
Incomplete outcome data	Low	Unclear	Low	Unclear	High
Selective outcome reporting	Low	Low	Low	Low	Low
Other potential threats to validity	Unclear	Unclear	Unclear	Unclear	Low
Summary assessment	**Low**	**Unclear**	**Low**	**Unclear**	**High**

5.3.1. Feeding characteristics and nutritional status

The two studies investigating the effect of PSIO appliances on feeding and nutritional status of cleft lip and/or palate (Prahl et al., 2005; Masarei et al., 2007) were considered overall as being at unclear risk of bias. This is because it is unclear whether the dropouts observed, although described and explained, could introduce some kind of bias; as well as, the effect of the varying compliance could have on their results. Regarding the domain selective outcome reporting, the studies were assessed to be of low risk of bias. Comments regarding the domains of random sequence generation, allocation concealment, blinding of participants and personnel, as well as, blinding of outcome assessment, have been provided previously.

5.3.2. Facial esthetics

The three studies investigating the effect of PSIO appliances on facial esthetics of cleft lip and/or palate children (Prahl et al., 2006; Bongaarts et al., 2008; Chang et al., 2014) were considered overall as being at low risk of bias. The problems observed in the incomplete outcome data and the other threats domains were not considered to constitute an overall risk of bias (Prahl et al., 2006; Bongaarts et al., 2008). Regarding the domain selective outcome reporting, the studies were assessed to be of low risk of bias. Comments regarding the domains of random sequence generation, allocation concealment, blinding of participants and personnel, as well as, blinding of outcome assessment, have been provided previously.

5.3.3. Dentofacial cephalometric variables

The Bongaarts et al. (2009) study investigating the effect of PSIO appliances on dentofacial cephalometric variables of cleft lip and/or palate children was considered overall as being at low risk of bias. The problems observed in the incomplete outcome data and the other threats domains

were not considered to constitute an overall risk of bias. Regarding the domain selective outcome reporting, the study was assessed to be of low risk of bias. Comments regarding the domains of random sequence generation, allocation concealment, blinding of participants and personnel, as well as, blinding of outcome assessment, have been provided previously.

5.3.4. Maxillary dentoalveolar variables

The three publications investigating the effect of PSIO appliances on maxillary dentoalveolar variables of cleft lip and/or palate children (Prahl et al., 2001; 2003; Bongaarts et al., 2006) were considered overall as being at low risk of bias. The problems observed in the incomplete outcome data in the Bongaarts et al. (2006) study and the other threats domains were not considered to constitute an overall risk of bias. Regarding the domain selective outcome reporting, the studies were assessed to be of low risk of bias. Comments regarding the domains of random sequence generation, allocation concealment, blinding of participants and personnel, as well as, blinding of outcome assessment, have been provided previously.

5.3.5. Dental arch relationships

The two studies investigating the effect of PSIO appliances on dental arch relationships of cleft lip and/or palate children (Bongaarts et al., 2004; Noverraz et al., 2015) were considered overall as being at low and unclear risk of bias. The problems observed in the incomplete outcome data in the Bongaarts et al. (2004) study and the other threats domains were not considered to constitute an overall risk of bias. Regarding the domain selective outcome reporting, the Noverraz et al. (2015) study was assessed to be at unclear risk of bias as they did not report on many outcomes, which were examined in the previous Bongaarts et al. (2004) investigation that had exactly the same research design. Comments regarding the domains of random sequence generation,

allocation concealment, blinding of participants and personnel, as well as, blinding of outcome assessment, have been provided previously.

5.3.6. Hearing, speech and language evaluation

All papers reporting on variables associated with hearing, speech and language evaluation (Konst et al., 1999; Konst 2000; 2002; 2003a; 2003b; 2003c) where considered to be at high risk of bias, mainly because of problems associated with incomplete outcome data. The problems observed in the other threats domain were not considered to constitute an overall risk of bias. Regarding the domain selective outcome reporting, the Konst et al. (2003a) was assessed to be at unclear risk of bias as they did not report on one outcome, which had been examined in the original report of the study in thesis form (Konst et al., 2002). Comments regarding the domains of random sequence generation, allocation concealment, blinding of participants and personnel, as well as, blinding of outcome assessment, have been provided previously.

5.3.7. Patient and caregiver-reported outcomes

The one study examining the effects of PSIO on caregiver-reported outcomes (Prahl et al., 2008) was assessed as being of unclear risk of bias, mainly because of problems associated with the incomplete outcome data and the other threats domains. Regarding the domain selective outcome reporting, the studies were assessed to be of low risk of bias. Comments regarding the domains of random sequence generation, allocation concealment, blinding of participants and personnel, as well as, blinding of outcome assessment, have been provided previously.

5.3.8. Economic evaluation related outcomes

From the four studies providing data on economic evaluation related outcomes, two from the DUTCHCLEFT sample (Severens et al., 1998; Konst et al., 2004) were considered to be at high risk, the Masarei et al. (2007) at unclear risk and the Chang et al. (2014) at low risk of bias. The first two studies had particular problems in the incomplete outcome data domain. For the same domain, the Masarei et al. (2007) paper received an unclear risk of bias assessment. The problems observed in the other threats domain were not considered to constitute an overall risk of bias (Konst et al., 2004; Masarei et al., 2007). Comments regarding the domains of random sequence generation, allocation concealment, blinding of participants and personnel, as well as, blinding of outcome assessment, have been provided previously.

5.3.9. Adverse effects and problems related to PSIO appliances and procedures

The Masarei et al. (2007) and the Chang et al. (2014) studies reporting on adverse effects and problems related to PSIO appliances and procedures, received assessments of being at unclear and low risk of bias respectively, for reasons already presented.

5.4. Effect of pre-surgical infant orthopedics vs. no treatment

The results of the studies included in the present review are presented below. Despite the limited available data, exploratory quantitative syntheses were conducted where possible. Because it was not possible to retrieve a sufficient number of trials, analyses for "small-study effects" and publication bias (Higgins and Green, 2011) were not able to be performed.

5.4.1. Feeding characteristics and nutritional status

The effect of pre-surgical infant orthopedics on feeding and nutritional status of cleft lip and/or palate subjects was investigated in two studies eligible for inclusion in the present review (Prahl et al., 2005; Masarei et al., 2007).

Prahl et al. (2005) from the DUTCHCLEFT sample evaluated the ***feeding*** records of in total 47 infants as soon as possible after birth and at 3, 6, 15, 24 weeks of age and observed no significant differences between the two groups regarding time per feeding (min), amount per feeding (mL) and feeding velocity (mL/min). The application of the PSIO appliance, cleft width and initial cleft width explained less than 4.0% of the variance observed in feeding characteristics outcomes. The Masarei et al. (2007) study investigated the effect of PSIO appliances in a total of 33 UCLP infants and 16 cleft palate infants and found that no participant from any group was rated as having a normal feeding pattern with the Neonatal Oral Motor Assessment Scale (NOMAS) prior to surgical closure of the soft palate, at three months of age. Similarly, no statistically significant differences were noted regarding physiological measures of bottle feeding and sucking (length of sucking bursts, peak-to-peak intervals, rate of sucking, suck-swallow rations and percentage pressure generation above baseline pressure in the feeding bottle) using the Great Ormond Street Measurement of Infant Feeding (GOSMIF). Finally, there was no significant difference between the intervention and control groups for either type of cleft for the number of abnormal swallowing behaviors as assessed by videofluoroscopy (21 patients). At twelve months of age, 6 months after surgical closure of the soft palate, all infants were rated as having normal oral motor skills using the Schedule of Oral Motor Assessment (SOMA) (21 patients).

Regarding ***nutritional status***, no statistically significant benefit was observed in weight (z scores) of PSIO treated UCLP children at 3-4 months, before any surgical intervention, [Weighted Mean Difference (WMD): 0.120, (95% Confidence Interval (CI): -0.338 - 0.578), p=0.607; 2 studies, n

= 72; I^2 = 0%] (Figure 2) and at around 12 months of age, after surgical closure of the soft palate [WMD: 0.132, 95% CI: -0.364 - 0.628, *p*=0.628; 2 studies, n = 42; I^2 = 0%] (Figure 2).

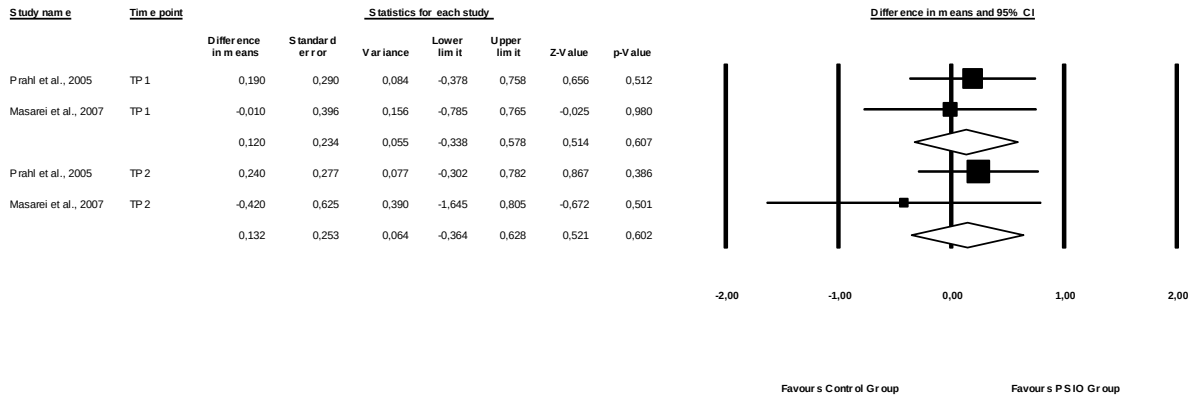


Figure 2. Weight (z scores) of PSIO treated UCLP children compared to control at 3-4 months, before any surgical intervention (TP1: Time Point 1), and at around 12 months of age, after surgical closure of the soft palate (TP2: Time Point 2).

Similarly, no statistically significant benefit was observed in height (z scores) of PSIO treated UCLP children at 3-4 months, before any surgical intervention, [WMD: -0.056, 95% CI: -0.535 - 0.424, *p*=0.820; 2 studies, n = 72; I^2 = 0%] (Figure 3) and at around 12 months of age, after surgical closure of the soft palate [WMD: -0.007, 95% CI: -0.502 - 0.489, *p*=0.979; 2 studies, n = 52; I^2 = 0%] (Figure 3). Moreover, Prahl et al. (2005) observed that the mean z scores for weight-for-length in the UCLP intervention group children were significantly lower after soft plate closure. Similarly, Masarei et al. (2007) noted that infants did not differ in head circumference at three months, or in head circumference and Body Mass Index at the twelve-month assessment. Finally, no differences were noted for the isolated cleft palate group, at any time point (16 and 11 patients, respectively).

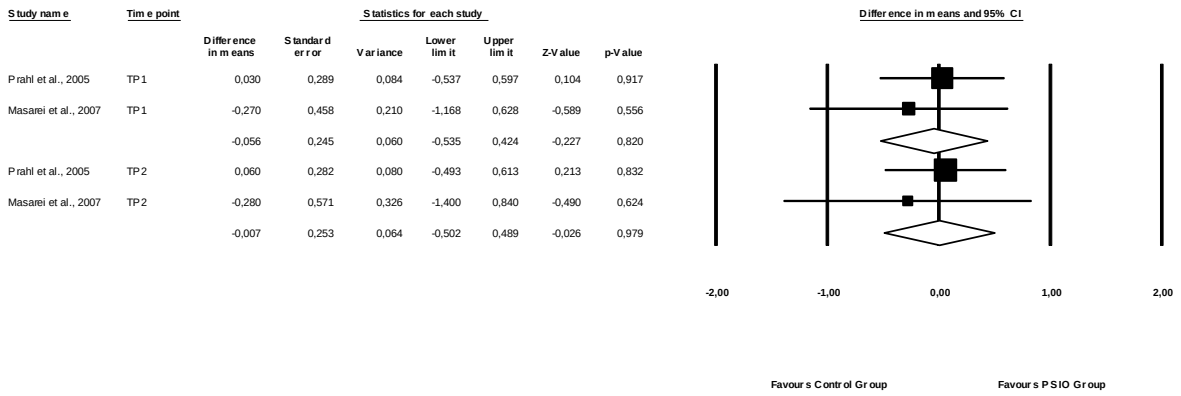


Figure 3. Height (z scores) of PSIO treated UCLP children compared to control at 3-4 months, before any surgical intervention (TP1: Time Point 1), and at around 12 months of age, after surgical closure of the soft palate (TP2: Time Point 2).

Based on these two papers (Prahl et al., 2005; Masarei et al., 2007), the quality of available evidence, using the GRADE approach (Guyatt et al., 2011), for the outcomes of weight and height evaluated at approximately 12 month of age was considered as very low (Table 9).

Table 9. Quality of available evidence for the outcomes of weight and height.

Quality assessment						№ of patients		Effect	Quality
Studies	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**PSIO**	**Control**	**Absolute (95% CI)**	
Weight [follow up: around 12 months; assessed with: z scores]									
2	Serious[1]	Not serious	Serious[2]	Serious[3]	None	27	25	MD **0.132 z scores higher** (0.364 lower to 0.628 higher) *p*=0.628	⊕⊕⊕◯ **VERY LOW**
Height [follow up: around 12 months; assessed with: z scores]									
2	Serious[1]	Not serious	Serious[2]	Serious[3]	None	27	25	MD **0.007 z scores lower** (0.502 lower to 0.489 higher) *p*=0.979	⊕⊕⊕◯ **VERY LOW**

CI: Confidence interval; MD: Mean difference

[1] Papers included were considered to be at unclear risk of bias. [2] Results were based on specific populations and treatment protocols. [3] The number of patients analyzed was limited.

5.4.2. Facial esthetics

Two studies analyzing data from the DUTCHCLEFT sample at various ages discussed the effects of PSIO on facial esthetics (Prahl et al., 2006; Bongaarts et al., 2008). In the first study, 41 subjects of 18 months of age, 21 of them treated with the PSIO appliance, were assessed, using photographs, by professional and adult laypeople judges. Two photographs, one showing the full face and a cropped version of the same photograph focusing on the mouth and nose area, were used to judge facial appearance by comparison to similar versions of a reference photograph (i.e. a good and clear photograph in the middle of the range from poor to excellent aesthetic outcome). The experimental photograph was presented together with the respective reference one and received a comparative score and a position on a Visual Analog Scale (VAS) and these scores were then pooled together. The score for the reference photograph was arbitrarily set a 100 and the assessors were instructed to increase it or decrease for the test view if they believed the esthetic outcome was better or worse respectively. Similarly, the reference photograph was given a specific position on the VAS and the judges were asked to move this position towards the right limit of the scale if they believed the outcome in the test photo was more esthetic compared to the reference one, or the opposite. No statistically significant differences were found between the intervention and the control group for the full face or the cropped photographs (z-scores).

In a subsequent study from the same sample (Bongaarts et al., 2008), 45 children (24 intervention and 21 control group) were evaluated at the age of 4 and 46 children (22 intervention and 24 control group) at the age of 6 years with a similar methodology. At the age of 4 years, the full face photos of subjects in the treatment group were judged to be more attractive than photos of children belonging to the control group. However, at 6 years of age this difference persisted only for the photos cropped to focus on the mouth and the nose judged by professionals. Regression analysis showed a minor effect of occlusion on the judgments of the full face photographs.

The quality of available evidence assessed with the GRADE was considered as low (Table 10).

Table 10. Quality of available evidence for facial esthetics assessment.

Quality assessment						№ of patients		Effect	Quality
Studies	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**PSIO**	**Control**	**Absolute (95% CI)**	
Full face photographs [follow up: 6 years of age; assessed with: laymen Visual Analog Scale scores compared to reference photograph]									
1	Not serious	Not serious	Serious[1]	Serious[2]	None	22	24	MD **4.520 scores higher** (-1.953 lower to 10.993 higher) *p*=0.080	⊕⊕◯◯ **LOW**
Cropped photographs [follow up: 6 years of age; assessed with: laymen Visual Analog Scale scores compared to reference photograph]									
1	Not serious	Not serious	Serious[1]	Serious[2]	None	22	24	MD **6.920 scores higher** (-1.281 lower to 15.121 higher) *p*=0.100	⊕⊕◯◯ **LOW**

CI: Confidence interval; MD: Mean difference

[1] Results were based on specific populations and treatment protocols. [2.] The number of patients analyzed was limited.

5.4.3. Dentofacial cephalometric variables

One study analyzing data from the DUTCHCLEFT sample investigated the effects of PSIO on dentofacial cephalometric variables (angular, linear and ratio variable representing soft and hard tissues, as well as, dental structures) (Bongaarts et al., 2009). Forty-one children (21 intervention and 20 control group) were evaluated at the age of 4 and 43 children (21 intervention and 22 control group) at the age of 6 years with a similar methodology. At first assessment, the interincisal angle was about 9 degrees larger in the treated group, a difference not verified in the latter evaluation. However, at 6 years of age, the mentolabial angle was almost 9 degrees smaller in the intervention than the control group.

The quality of the available evidence for selected cephalometric measurements assessed with the GRADE approach was considered as low (Table 11).

Table 11. Quality of available evidence for selected cephalometric measurements [SNA & ANB angles].

Quality assessment						№ of patients		Effect	Quality
Studies	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**PSIO**	**Control**	**Absolute (95% CI)**	
SNA [follow up: 6 years of age; assessed with: degrees (º)]									
1	Not serious	Not serious	Serious[1]	Serious[2]	None	21	20	MD **1.290 degrees (º) lower** (-3.858 lower to 1.278 higher) *p*=0.306	⊕⊕◯◯ **LOW**
ANB [follow up: 6 years of age; assessed with: degrees (º)]									
1	Not serious	Not serious	Serious[1]	Serious[2]	None	15	16	MD **1.000 degrees (º) lower** (-3.201 lower to 1.201 higher) *p*=0.362	⊕⊕◯◯ **LOW**

CI: Confidence interval; MD: Mean difference

[1] Results were based on specific populations and treatment protocols. [2.] The number of patients analyzed was limited.

5.4.4. Maxillary arch dentoalveolar variables

Three publications analyzing data from the DUTCHCLEFT sample investigated the effects of PSIO on maxillary dentoalveolar variables at different ages (Prahl et al., 2001; 2003; Bongaarts et al., 2006).

Prahl et al. (2003) and Bongaarts et al. (2006) studied the ***contact and collapse of the alveolar segments*** in maxillary casts at various ages. In the first study, a total of 49 children (24 in the intervention group and 25 in the control) were assessed shortly after birth when comparable arch forms with no contact or overlap of the maxillary segments were observed in both groups. At subsequent assessments at 15 weeks of age (prior to lip repair), 24 weeks (6 weeks after lip repair), 48 weeks (4 weeks prior to soft palate closure), 58 weeks (approximately 6 weeks after soft palate surgery and at 78 weeks (half a year after soft palate closure) the frequency of collapse and its severity increased similarly between the two groups. Bongaarts et al. (2006) produced corroborating evidence at the 4 year and 6 year assessments, after examining the casts of 45 children (23 in the intervention group and 22 in the control at the first measurement, and 22 in the intervention group and 23 in the control at the second).

Various ***linear and angular maxillary dentoalveolar variables*** changes were assessed three dimensionally in casts and presented in the papers of Prahl et al. (2001) (24 children in the intervention and 25 in the control group; assessed shortly after birth and at 15 weeks (before lip surgery), 24 weeks, 48 weeks, 58 weeks and 78 weeks of age (6 months after soft palate surgery)) and Bongaarts et al. (2006) (4 years assessment: 23 children in the intervention and 22 in the control group; 6 years assessment: 22 children in the intervention and 23 in the control group).

At the 78 week examination, the only statistically significant difference, between the two groups in the assessed variables, was observed in the change from the baseline of anterior arch depth. This variable increased more in the intervention group compared to the control children. No other statistically significant change from the baseline was noted in variables regarding cleft width (alveolar cleft width, midpalatal cleft width, posterior cleft width at the tuberosity level), as well as, linear arch dimensions (anterior arch width, tuberosity width, total arch depth, total arch length, alveolar cleft margins length) and angular measurements considered (angulation of the greater alveolar segment in the transverse plane of space, angulation of the smaller alveolar segment in the transverse plane of space, vertical slope of the greater alveolar segment, vertical slope of the smaller alveolar segment).

The comparison of the various maxillary arch variables at 4 years of age revealed a statistically significant increase in total arch depth in the treated group and the angulation of the greater alveolar segment in the transverse plane of space. At the 6-year re-evaluation, no statistically significant difference was observed between the two experimental groups.

The quality of the available evidence for selected variables was considered as low (Table 12).

Table 12. Quality of available evidence for selected maxillary arch dentoalveolar variables.

Quality assessment						№ of patients		Effect	Quality
Studies	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**PSIO**	**Control**	**Absolute (95% CI)**	
Arch width (P2-P2') [follow up: 6 years of age; assessed with: mm]									
1	Not serious	Not serious	Serious[1]	Serious[2]	None	22	22	MD **0.980 mm lower** (-2.641 lower to 0.681 higher) *p*=0.240	⊕⊕◯◯ **LOW**
Arch depth (I-TT') [follow up: 6 years of age; assessed with: mm]									
1	Not serious	Not serious	Serious[1]	Serious[2]	None	15	16	MD **0.110 mm higher** (-1.855 lower to 2.075 higher) *p*=0.910	⊕⊕◯◯ **LOW**

CI: Confidence interval; MD: Mean difference

[1] Results were based on specific populations and treatment protocols. [2.] The number of patients analyzed was limited.

5.4.5. Dental arch relationships

The effect of PSIO on dental arch relationships was investigated in two papers analyzing data from the DUTCHCLEFT sample at different time points (Bongaarts et al., 2004; Noverraz et al., 2015). Bongaarts et al. (2004) evaluated occlusion at 4 years (22 children in the intervention and 22 in the control group) and at 6 years (22 children in the intervention and 23 in the control group). No statistically significant differences were observed regarding the 5-year old index (categorizing arch relationships using reference models), the Huddart/Bodenham score (dental arch relationships at the transverse planes), overjet, overbite and sagittal occlusion assessment (scored for deciduous canines and second deciduous molars according to Angle classification). Corroborating evidence was produced by assessing occlusion with the Huddart/Bodenham score modified for mixed dentition at 9 (24 children in the intervention and 21 in the control group) and 12 (22 children in the intervention and 22 in the control group) years of age (Noverraz et al., 2015).

The quality of the available evidence for selected variables was considered as low (Table 13).

Table 13. Quality of available evidence for selected Huddart/Bodenham scores.

Quality assessment						№ of patients		Effect	Quality
Studies	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**PSIO**	**Control**	**Absolute (95% CI)**	
Total Huddart/Bodenham score [follow up: 12 years of age; assessed with: points]									
1	Not serious	Not serious	Serious[1]	Serious[2]	None	22	22	MD **0.510 points lower** (-3.57 lower to 2.54 higher) *p*=0.736	⊕⊕◯◯ **LOW**
Buccal Huddart/Bodenham score for the cleft side [follow up: 12 years of age; assessed with: points]									
1	Not serious	Not serious	Serious[1]	Serious[2]	None	22	22	MD **0.030 points lower** (-1.700 lower to 1.630 higher) *p*=0.977	⊕⊕◯◯ **LOW**

CI: Confidence interval; MD: Mean difference

[1] Results were based on specific populations and treatment protocols. [2.] The number of patients analyzed was limited.

5.4.6. Hearing, speech and language evaluation

The evaluation of different aspects of speech and language development in the DUTCHCLEFT sample of patients started at age 1 and continued at 6-month-intervals until 3 years of age. Language development was also assessed at a 6-year follow-up. Subjects included had both parents fluent in the Dutch language and were first evaluated regarding pre-lexical development (sounds produced in babbling) at age 1 and 1.5 (Konst et al., 1999). Subsequently, phonological development of the children was followed from age 2 to 3 (Konst et al., 2003a) and an interim speech quality and intelligibility assessment was performed at age 2.5 (Konst et al., 2000; 2002; 2003c). Moreover, during the same time period (from age 2 to 3), patients' hearing status and language skills development was investigated. Finally, expressive language skills were re-assessed at six years of age (Konst et al., 2003b).

Assessment of ***pre-lexical development*** (Koopmans-van Beinum and Van de Stelt's system: classification of utterances, phonation characteristics, articulation movement, phonetic inventory) was performed in 36 infants aged 12 months (18 in the intervention and 18 in the control group) and 38 infants aged 18 months (19 in the intervention and 19 in the control group). Development from 12 to 18 months was followed in 31 children (15 in the intervention and 16 in the control group). At 12 months of age, babies using the PSIO appliance showed improved use of alveolar

articulations, however, the effects of PSIO seemed to be of a transient nature. At age 1.5 years, changes towards more normal articulation occurred in both groups and no significant differences in all variables of sound production were noted between them (Konst et al., 1999).

Subsequent ***phonological development*** was investigated with a system specific for Dutch children (Fonologische Analyse van het Netherlands (FAN): number of acquired consonants and order of phonological development, use of phonological processes, occurrence of nasal escape) at age 2, 2.5 and 3 years (Konst et al., 2003a). Only six toddlers (4 in the intervention and 2 in the control group) were assessed at all three time points. At the 2-year evaluation, 16 children participated (9 in the intervention and 7 in the control group), at the 2.5-years group, 18 children were assessed (9 in the intervention and 7 in the control group), and at 3 years of age, 12 toddlers (6 in the intervention and 7 in the control group) were evaluated. Subjects from the intervention group exhibited a more normal path of phonological development during the observation period. At the 2.5 year evaluation, this was classified as normal or delayed, whereas most control toddlers were categorized as belonging to the abnormal developmental pattern. In addition, at the age 3 assessment, children of the former group had acquired more initial consonants than those in the non-treated one. No other statistically significant differences were observed in the use of phonological processes or the occurrence of nasal escape.

During the same chronological age interval, the ***speech quality*** of 10 toddlers 2.5 years-old (10 in the intervention and 10 in the control group) was evaluated perceptually (on seven-point equal-appearing interval rating scales) regarding 13 specific characteristics (*place of articulation*: palatalization, lateralization, fronting, backing, glottal articulation; *voice characteristics*: hyperkinetic voice; *nasalization*: hypernasality, nasal emission, nasal fricative, nasal snort, nasal realization; *general evaluative characteristics*: correctness of articulation, intelligibility) together with an assessment of an overall impression of speech quality (Konst et al., 2003c) and of the need for speech therapy in the year following the experiment (Konst, 2002). Based on the results of this

study, children belonging to the intervention group obtained significantly higher ratings for intelligibility than the control group children. This result was in agreement with the finding of an earlier study (10 toddlers 2.5 years-old; 10 in the intervention and 10 in the control group), which had assessed speech intelligibility by means of rating on a ten-point rating scale marked by the contrasting labels 'unintelligible' (rating 1) to 'intelligible' (rating 10) (Konst et al., 2000). However, these ratings did not completely reflect intelligibility defined as the proportion of words understood by the listener, as data obtained by means of transcriptions indicated that, in fact, there were no group differences in actual intelligibility.

The evaluation of different aspects of speech and language in the DUTCHCLEFT sample was complemented by assessment of patients' ***language skills development*** (Konst et al., 2003b). At the age of 2, 2.5, and 3 years linguistic development was evaluated in a group of 12 toddlers (6 in the intervention and 6 in the control group). Receptive language skills were investigated using the Dutch version of the Reynell Developmental Language Scales and expressive language skills by calculating the variables mean length of utterance (MLU) and mean length of longest utterances (MLLU). No differences in receptive language skills were noted, however, at age 2.5 and 3 years, the toddlers belonging to the treated group performed better regarding expressive language skills. In the fraction of these patients (6 in the intervention and 5 in the control group), which were re-evaluated at 6 years of age using standardized Dutch language tests (expressive vocabulary: Taaltests voor Kinderen; expressive syntactic skills: Schlichting test), the difference in expressive language between the experimental groups ceased to be significant.

Evaluation of the ***hearing status*** in the context of the latter study (Konst et al., 2003b) revealed that hearing thresholds and presence of middle early infection did not present gross differences between the two groups at the age of 2, 2.5, and 3 years.

The quality of the available evidence for speech intelligibility at 2.5 year-old toddlers was considered as very low (Table 14).

Table 14. Quality of available evidence for speech intelligibility.

Quality assessment						№ of patients		Effect	Quality
Studies	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**PSIO**	**Control**	**Absolute (95% CI)**	
Speech intelligibility score [follow up: 2.5 years of age; assessed with: points]									
1	Serious[1]	Not serious	Serious[2]	Serious[3]	None	10	10	MD **1.180 points higher** (0.202 lower to 2.157 higher) *p*=0.020	⊕⊕⊕◯ **VERY LOW**

CI: Confidence interval; MD: Mean difference

[1] The study was considered of high risk of bias. [2]Results were based on specific populations and treatment protocols. [3.] The number of patients analyzed was limited.

5.4.7. Patient and caregiver-reported outcomes

One publication presenting data from the DUTCHCLEFT sample investigated the effects of PSIO on caregiver-reported outcomes (Prahl et al., 2008).

In total, the mothers of 54 infants were asked to complete and return a questionnaire when their children were 6, 24 and 58 weeks old. The questionnaire contained 42 questions, completed on a 4-point scale (from 1 - very satisfactory, very happy, a lot of fun, more than adequate - to 4 - very unsatisfactory, very unhappy, no fun, very inadequate) and categorized in four domains.

Domain 1: Interaction and Caretaking of the Baby (pleasure experienced during interaction when cuddling, consoling, playing, walking, visiting, and caretaking when feeding, bathing, dressing, and changing, including interaction when baby was crying and getting the baby out of bed).

Domain 2: Comings and Goings of the Baby (feelings of the mother toward the comings and goings of the baby, i.e. sleeping, feeding, and the timing of both, digestion, conduct during the day, and reaction of the baby during interaction and when left to play by himself or herself).

Domain 3: Motherhood and Life Outside (feelings of the mother toward motherhood - specifically and in terms of quality of life - and life outside of motherhood - i.e. relationship with partner, own spirit, family spirit, and housekeeping, sufficient time for other occupations or pursuits than the baby within the house or outside the house, time for self and time for friends and relatives).

Domain 4: Support (support from partner, relatives, acquaintances and friends, and others for mental support, caretaking, housekeeping; support, information, and advice from the cleft team, general physician, and other medical institutions; adequate contact with other parents and parents of children with clefts).

Statistical analysis of the questionnaires finally provided by 49 mothers did not locate any significant differences between the two groups. The quality of the available evidence for the mean satisfaction scores of the derived from the questionnaire as a whole (four domains) was considered as low (Table 15).

Table 15. Quality of available evidence for total questionnaire scores.

Quality assessment						№ of patients		Effect	Quality
Studies	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**PSIO**	**Control**	**Absolute (95% CI)**	
Total questionnaire score [follow up: approximately 1 year of age; assessed with: points]									
1	Not serious	Not serious	Serious[1]	Serious[2]	None	18	19	MD **0.100 points higher** (-0.05 lower to 0.24 higher) *p*=0.352	⊕⊕◯◯ **LOW**

CI: Confidence interval; MD: Mean difference

[1] Results were based on specific populations and treatment protocols. [2,] The number of patients analyzed was limited.

5.4.8. Economic evaluation related outcomes

Three papers eligible for inclusion in the present thesis gave comparison data on economic evaluation related outcomes, two of them from the DUTCHCLEFT sample (Severens et al., 1998; Konst et al., 2004; Masarei et al., 2007).

Regarding analysis of the ***costs associated with PSIO appliance treatment and calculations on cost-effectiveness***, until surgical closure of the lip (around 18 weeks of age) total direct medical costs (calculated on the cost of personnel, cost of PSIO materials and overhead cost) expressed in US$ (1994 base year) in the intervention group were significantly more (852 ±69 US$ for 23 infants) that in the control group (304 ±34 US$ for 20 infants). The same was observed for indirect non-medical costs (calculated on the number of visits, the number of parents that would normally

accompany the child, the average time of the visit, the duration of travel, and the normal daily activity of the accompanying parent (s)) expressed in 1994 base year US$ (231 ±106 US$ for 15 infants in the intervention group and 130 ±14 US$ for 14 infants in the control group). However, no statistically significant difference was observed between two groups with regards to the duration of surgical lip closure (23 children in the intervention and 20 children in the control group) (Severens et al., 1998). Further details on direct medical and travel costs from the Severens et al. (1998) data are presented in Table 16.

Table 16. Data on medical and travel costs expressed in US$ (1994 base year) in the PSIO treated and the control group until surgical closure of the lip, around 18 weeks of age (adapted from Severens et al., 1998).

	Intervention group			Control group		
Domain	**No.**	**Mean**	**SD**	**No.**	**Mean**	**SD**
Medical costs						
Cost of personnel	23	191	61	20	108	34
Cost of PSIO materials	24	175	56	24	0	0
Overhead cost	24	479	69	22	196	34
Non-medical costs						
Travel cost	23	128	106	14	130	44

For the whole duration of PSIO appliance treatment, the costs submitted by the orthodontist in € (2002 base year) were 1460 ±247 € for 10 infants in the intervention and 419 ±91 € for 10 infants in the control group (Konst et al., 2004). The calculated cost effectiveness from the perspective of speech development (total impression of speech quality at the age of 2.5 years) was 1041 € for 1.34 points of speech quality improvement.

The quality of the available evidence for the cost of treatment by the orthodontist in the period from birth to soft palate closure was considered as very low (Table 17).

Table 17. Quality of available evidence for total cost of treatment by the orthodontist.

Quality assessment						№ of patients		Effect	Quality
Studies	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**PSIO**	**Control**	**Absolute**	
Speech intelligibility score [follow up: 2.5 years of age; assessed with: points]									
1	Serious[1]	Not serious	Serious[2]	Serious[3]	None	10	10	MD **1041 € higher**	⊕⊕⊕◯ **VERY LOW**

MD: Mean difference

[1] The study was considered of high risk of bias. [2]Results were based on specific populations and treatment protocols. [3.] The number of patients analyzed was limited.

Regarding the ***visits to the orthodontist***, for the first 18 weeks, 7.2 ±1.8 were needed for the 23 infants in the intervention and 2.4 ±0.6 € for 20 infants in the control group (Severens et al., 1998). For the whole period of PSIO appliance intervention, statistical significantly more visits were reported for the PSIO group, both in the Prahl et al. (2001) study (for the 12 month duration of PSIO treatment; mean difference: 9.4; 95% CI: 8.02-10.78) and the Masarei et al. (2007) study (for the 6 month duration of PSIO treatment; mean difference: 2.68; 95% CI: 0.83-4.53)

5.4.9. Adverse effects and problems related to PSIO appliances and procedures

Possible adverse effects and problems related to PSIO appliances and procedures were not systematically investigated in the studies included in the present review comparing the effect of PSIO appliances to no treatment.

Only the Masarei et al. (2007) paper reported on possible problems related to PSIO appliances and procedures. In approximately 20 percent of infants (9 out of 50) the operators reported loose fitting plates, which were corrected. Other minor problems included oral thrush, minor ulceration and neonatal teeth.

5.5. Comparative effect of PSIO procedures

Only one study (Chang et al., 2014) compared the effects between two PSIO procedures. The results are presented below.

5.5.1. Facial esthetics

No difference was noted for nostril height, nostril sill height and nostril area ratio at any available time point of measurement. Only for the variable nostril width ratio was an increase observed after PSIO procedures with the Figueroa technique. However, six months after surgical correction this difference had ceased to exist.

The quality of the available evidence for nostril height and nostril width ratios was considered as low (Table 18).

Table 18. Quality of available evidence for selected variables.

Quality assessment						№ of patients		Effect	Quality
Studies	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**mF**	**mG**	**Absolute (95% CI)**	
Nostril Height ratio [follow up: approximately 9 months of age; assessed with: points]									
1	Not serious	Not serious	Serious[1]	Serious[2]	None	15	15	MD **0.000 points** (-0.097 lower to 0.097 higher) *p*=0.934	⊕⊕◯◯ **LOW**
Nostril width ratio [follow up: approximately 9 months of age; assessed with: points]									
1	Not serious	Not serious	Serious[1]	Serious[2]	None	15	15	MD **0.286 points higher** (-0.203 lower to 0.663 higher) *p*=0.313	⊕⊕◯◯ **LOW**

mF: modified Figueroa technique; mG: modified Grayson technique; CI: Confidence interval; MD: Mean difference

[1] Results were based on specific populations and treatment protocols. [2.] The number of patients analyzed was limited.

5.5.2. Economic evaluation related outcomes

Regarding analysis of the ***costs associated with PSIO appliance treatment***, no statistically significant difference was noted with regards to total costs of treatment for parents/caregiver (Figueroa group 1240 ±250.14 US$; Grayson group 1159 ±275.81 US$; *p*=0.357) and total cost

of treatment for national insurance (Figueroa group 20267 ±1668 US$; Grayson group 19667 ±1839 US$; *p*=0.357).

The quality of the available evidence for the costs associated with PSIO appliance treatment was considered as low (Table 19).

Table 19. Quality of available evidence for selected variables.

Quality assessment						№ of patients		Effect	Quality
Studies	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**mF**	**mG**	**Absolute (95% CI)**	
Total costs of treatment for parents/caregiver [follow up: approximately 9 months of age; assessed with: points]									
1	Not serious	Not serious	Serious[1]	Serious[2]	None	15	15	MD **81.00 US$ higher** (-115.93 lower to 277.93 higher) *p*=0.357	⊕⊕◯◯ **LOW**
Total cost of treatment for national insurance [follow up: approximately 9 months of age; assessed with: points]									
1	Not serious	Not serious	Serious[1]	Serious[2]	None	15	15	MD **600.00 US$ higher** (-713.1 lower to 1913.1 higher) *p*=0.357	⊕⊕◯◯ **LOW**

mF: modified Figueroa technique; mG: modified Grayson technique; CI: Confidence interval; MD: Mean difference
[1] Results were based on specific populations and treatment protocols. [2.] The number of patients analyzed was limited.

Regarding the ***visits to the orthodontist(s)***, no statistically significant difference was observed regarding the total number of visits before surgical correction (Figueroa group 8.33 ±1.59; Grayson group 7.67 ±1.84; *p*=0.297).

5.5.3. Adverse effects and problems related to PSIO appliances and procedures

Regarding adverse effects and problems related to PSIO appliances and procedures, only a statistically significantly increased occurrence of alveolar ulceration in the Grayson group was noted. No other differences were observed regarding nasal ulceration, skin rash or superficial skin injury caused by taping.

6. DISCUSSION

Clefts of lip, alveolar ridge and palate are considered very common congenital malformations in humans (Mossey and Little, 2002). These defects involve, to varying degrees and extent, soft and hard tissue components of the oro-facial region and may not only affect many basic human activities, such as feeding and speech, but also influence other important domains like esthetic self-perception, as well as satisfaction and quality of life of the patient itself and its entourage (Parhl et al., 2008). Thus, this complex multifactorial anomaly can have a huge impact on the life of both patients and caregivers physiologically, psychologically, socially, not disregarding economically. The cleft palate team's main challenge and aim is to deliver effective care without side effects, so as to improve the quality of life of their patients (Prahl-Andersen and Ju, 2006, Shaw, 2000). Moreover, as resources and professional help are always limited in the context of health-care systems, treatment must be optimally cost effective (Severens et al., 1998).

PSIO concept has been integrated into the standards of care for cleft lip and/or palate patients in many dedicated treatment teams around the world (Shaw et al., 1992; Prahl et al., 2001; Lohmander et al., 2004; Masarei et al., 2007; Long et al., 2011). However, the relevant treatment outcomes have been a matter of controversy (Cash, 2012). Two systematic reviews evaluating PSIO in general (Uzel and Alparslan, 2011; Papadopoulos et a., 2012), and on assessing growth and development outcomes after various feeding interventions in infants with cleft lip and/or palate (Bessell et al., 2011) have been published recently. However, neither of these attempted to summarize the available quality of evidence and thus provide an insight into the strength of the relevant recommendations.

6.1. Summary of available evidence

From the initially identified records, twenty full-text reports from randomized controlled trials on children with unilateral cleft lip and palate, as well as, cleft palate were included in the systematic review. No studies on children with bilateral cleft lip and palate were retrieved.

Eighteen of the eligible publications reported on various outcomes originating from the sample of a single study. In essence, only three randomized controlled trials satisfied the pre-specified criteria for inclusion in the present paper, reflecting the scarcity of relevant research at the top of the widely accepted hierarchy of scientific evidence. Considerable research has been done in trials that were not randomized despite the fact that it is widely accepted that well-designed and properly executed RCTs provide the best evidence on the efficacy of health care interventions (Altman et al., 2001; Oxford Centre for Evidence-based Medicine, 2009). Among these three acceptable RCTs, only one (Dutchcleft) followed patients from approximately 2 weeks after birth until 12 years of age. The consequent lack of extensive data of high evidence based potential is rather surprising bearing in mind not only the prevalence of the problem (Mossey and Little, 2002), but also the fact that its manifestations are usually life-long and associated with significant morbidity (Cobourne, 2012). Thus, relevant, evidence-based information would be beneficial in order to support the care provided for this children with possible consequences for a broad spectrum of human life, like growth and development, esthetics, speech and language development, as well as, various psychosocial parameters.

6.1.1. Effect of pre-surgical infant orthopedics vs. no treatment

In general, based on the information provided from in the two study samples eligible for inclusion in the present review, PSIO does not exhibit clinically significant effects over no treatment.

Two papers (Prahl et al., 2005; Masarei et al., 2007) included outcomes related to **feeding and general body growth**; functions and outcomes considered of the utmost importance in evaluating

the early stages of human life. According to the results of the two studies, which were consistent with each other, neither passive nor active PSIO appliances, have discernible positive effects on feeding functioning and the subsequent nutritional status in patients with unilateral cleft lip and palate or isolated cleft palate. A recent systematic review of randomized controlled trials focusing exclusively on growth and development outcomes assessment after various feeding interventions in infants with cleft lip and/or palate reported similar findings (Bessell et al., 2011).

Furthermore, passive PSIO appliances did not generally exert a significant influence on **facial esthetic perception**, which constitutes another important concern related to this condition (Prahl et al., 2006; Bongaarts et al., 2008). At the longest follow-up at 6 years of age, the PSIO treated group got better ratings when professionals assessed photographs focusing on the nasolabial region. Judgments on cropped photographs are considered more reliable because they may help blind other parameters, such as general characteristics of the baby face and variation in the response to facial expressions that have been found to positively affect overall ratings (Hilderbrandt and Fitzgerald, 1978; Tobiasen, 1987; 1989; Tobiasen et al., 1991; Asher-McDade et al., 1991). However, this beneficial effect observed in professional ratings was considered irrelevant in daily life, where affected children interact with ordinary people.

Cephalometric investigation of **facial growth** in the Dutchcleft sample did not reveal any clinically relevant effect of passive PSIO appliances either (Bongaarts et al., 2009), although it is not totally clear whether their results could have been influenced by the uncertainty inherent in landmark localization (Baumrind and Frantz, 1971; Houston, 1983; Marci and Athanasiou, 1995; Trpkova et al., 1997). In addition to common errors in cephalometric investigations, the presence of malpositioned incisors in an atrophic and displaced premaxilla could be regarded as an additional cause of measurement errors in young cleft patients (Atherton, 1967; Bongaarts et al., 2008). In the case of unilateral cleft lip and palate toddlers, alternatives for cephalometric radiographics points A, ANS, and PNS have been assessed but not proven to perform better than traditional

landmarks (Bongaarts et al., 2008b). Overall, the level of measurement errors in the Bongaarts et al (2009) study was considered acceptable.

Investigation of orofacial region growth by means of study casts also failed to reveal any discernible effects of passive PSIO appliances at the longest follow-ups available. PSIO processes were not able to prevent collapse of the maxillary arch or influence maxillary arch dimensions in 6-year-old children (Bongaart et al., 2006). It may be worth noting that non-RCTs have produced data on significant differences for some of the maxillary arch measurements in younger children (Mishima 1996a; 1996b; 2000; 2001).

Moreover, assessment of occlusal relationships using the modification of the Huddart/Bodenham system showed no differences between treatment and control groups at 12 years of age (Noverraz et al., 2015), which conforms with the results of a non-RCT evaluating the effect of an active PSIO appliance with the GOSLON yardstick in younger children (Chan et al., 2003). This modification of the Huddart/Bodenham system used to examine dental arch relationships in the transverse place of space is regarded to be a powerful, valid and reliable measurement of treatment outcome in cleft lip and palate children (Mossey et al., 2003; Gray and Mossey, 2005; Altalibi et al., 2013).

In general, when considering overall growth in cleft lip and palate patients, it should be noted that since individualization of the treatment is expected, the type and frequency of the various therapeutic interventions may constitute confounding factors; something that remains to be investigated in depth (Noverraz et al., 2015). Moreover, growth and treatment parameters may not be the only factors affecting the general esthetic perception in patients with cleft lip and palate, a subject of particular and life-long importance (Bongaarts et al., 2008; 2009). Additional parameters, including, but not limited to, facial expression, skin texture, hair and eye color may be exert greater influence than expected, effects needing to be studied at long follow-ups as they may become more prominent following the pubertal growth (Bongaarts et al., 2008).

Regarding **speech and language development**, 2.5 year-old toddlers belonging to the intervention group not only obtained significantly higher ratings for intelligibility than the control group children (Konst et al., 2000), but these scores were high enough for the whole intervention to be considered cost-effective from a societal economic evaluation point of view (Konst et al., 2004). However, these results should be considered with caution; firstly because these ratings did not completely reflect intelligibility defined as the proportion of words understood by the listener (Konst et al., 2000) and secondly because no sound longer-term data on the broader subject exist. Only the data from a preliminary assessment of 12 patients' language skills development at 6 years of age suggested there was no difference between treated and untreated groups (Konst et al., 2003b). However, a similar lack of difference between the two groups was also noted in two relevant non-RCTs (Karling et al., 1993; Lohmander et al., 2004).

Finally, no statistically significant outcomes were observed regarding caregiver **psychosocial related outcomes** like satisfaction in motherhood (Prahl et al., 2008). However, these results were obtained using an instrument not previously tested and validated regarding its psychometric properties.

6.1.2. Comparative effect of PSIO procedures

No statistically important differences were observed between the modified Figueroa and the modified Grayson nasoalveolar molding techniques regarding nasal **facial esthetics** six months after surgical correction of the lip (Chang et al., 2014). Even though some studies have indicated an advantageous use of pre-surgical nasoalveolar molding for the improvement of nasal symmetry (Grayson et al., 1999; Grayson and Cutting, 2001; Grayson and Maull, 2004; Barillas et al., 2009; Gomez et al., 2012; Liao et al., 2012), these results are not universal (Ross and MacNamera, 1994) and remain to be supported by RCTs.

Both approaches, from an **economic evaluation outcomes** approach, incurred significant costs for parents/caregivers and national insurance institutions. The only difference noted was in **adverse effects and problems related to PSIO appliances and procedures**, namely a statistically significant increased occurrence of alveolar ulceration in the Grayson group.

6.2. Quality of the available evidence

Overall, the quality of evidence assessed with the GRADE approach (Guyatt et al., 2011) was considered at best as low, indicating caution regarding the strength of the relevant recommendations.

Available information on the vast majority of outcomes considered was based on data from one randomized controlled trial, namely the Dutchcleft study sample, indicating the **scarcity of evidence based information** on a problem with such significant and life-long consequences for both patients and their families. Only regarding general body growth outcomes during the first year of life, was the literature search able to retrieve data from two separate data sets and an exploratory quantitative data synthesis was attempted. The results obtained from these studies were similar and the I^2 statistic obtained from the meta-analytic calculations suggested a relatively unimportant degree of heterogeneity, indicating that **inconsistency** during the GRADE assessment was not to be considered serious. In the context of the present review heterogeneity can arise from diversity in terms of the characteristics of both patients and interventions, as well as measurement methodology related factors and was incorporated into a justifiable random effects model.

An important factor leading to downgrading the overall quality of evidence originated from the **risk of bias assessment** for the outcomes considered in the included studies. Overall, outcomes included in six of the retrieved publications studies were considered as being of low risk of bias (Prahl et al., 2001; 2003; Bongaarts et al., 2004; 2006; Prahl et al., 2006; Bongaarts et al., 2008; 2009; Chang et al., 2014) and four of unclear (Prahl et al., 2005; Masarei et al., 2007; Prahl et al.,

2008; Noverraz et al., 2015). The outcomes included in eight publications, regarding speech and language development and economic evaluation outcomes, were considered as being of high risk of bias (Severens et al., 1998; Konst et al., 1999; Konst, 2000; Konst et al., 2002; 2003a; 2003b; 2003c; 2004). As all studies included in the present review were considered to present low risk of bias regarding the domains of random sequence generation, allocation concealment, blinding of the participants, caregivers and the personnel and blinding of outcome assessment, the examination of the rest of the domains considered produced results that contributed to the variable overall assessment.

Many problems regarding risk of bias were attributable to incomplete outcome data. In two out of three patient samples included in the present review, patients finally analyzed were less than those originally randomized (Ducthcleft sample) or those initially calculated during power sample calculations (Masarei et al., 2007). Of course, it is absolutely expectable that in randomized controlled trials, especially those evaluating outcomes after lengthy time periods, a significant proportion of participants may be not evaluated for various reasons. Especially in the Ducthcleft study, an attempt to repeat the power calculations (based on SNA difference at 4 years of age) during study progression for some of the variables (SNA, ANB, 5-year index, esthetic score; Bongaarts et al., 2009) was made and it was the considered that intervention and control groups were large enough to show significant differences, if any were present. However, these calculations may not be applicable for other outcomes. Moreover, *post hoc* power calculations, with the G*Power software (version 3.1.9.2; Faul et al., 2007; 2009), using the actual results reported showed that the achieved power in most instances failed by a wide margin to achieve the initial 80% setting. Thus, in some cases the risk of bias from missing data was considered unclear while in others, especially those evaluating speech and language development and economic evaluation outcomes, the risk of bias regarding this domain and overall was considered high.

Some assessments of outcomes included in the material retrieved, were considered to present other

threats to validity. For example, in some cases, no definite conclusions could be reached regarding the effect that the varying compliance of PSIO appliances wear could have on study results. All included publications reported some relevant data, most probably derived from questioning the parents or the caregivers. However, this method is considered to be less effective despite its frequent use (Meyer-Gutknecht et al., 2014). In addition, in many instances the effect of various surgical or orthodontic interventions, supplementary to the original protocol, especially in the papers reporting data on the longer follow-ups was not clear-cut, thus creating uncertainties regarding possible bias in the retrieved data.

In all the cases considered, the overall quality of evidence was downgraded because of problems related to serious **indirectness** of the evidence retrieved and problems related to **imprecision**. The results obtained were derived from specific populations (particular ethnic background and cleft type) and treatment protocols; hence even this limited set of data cannot be applied with certainty in clinical settings characterized by a different patient mix or variable type and frequency of treatment modalities. Moreover, for varying reasons, the numbers of patients analyzed were limited, creating serious problems regarding the precision of the results obtained.

6.3. Strengths and limitations of the present review

The strengths of the present review include using a methodology following well-established guidelines and the fact that it focused exclusively on randomized controlled trials, as it is widely accepted that well-designed and properly executed RCTs provide the best evidence, with decreased risk of bias, on the efficacy of health care interventions (Altman et al., 2001; Oxford Centre for Evidence-based Medicine, 2009). The available empirical evidence suggests that intervention effects in orthodontic research seem to differ in non-RCTs compared to RCTs (Papageorgiou et al., 2015). The two systematic reviews published recently that evaluated PSIO appliances in general (Uzel and Alparslan, 2011; Papadopoulos et a., 2012), also included non-

RCTs. Furthermore, neither attempted to summarize the quality of available evidence and thus provide an insight into the strength of the relevant recommendations, as was carried out in the present systematic review based in the GRADE approach (Guyatt et al., 2011).

Moreover, the search strategy employed in the present review was both exhaustive, covering electronic, manual, and gray literature material up to May 2016, and comprehensive including every available randomized controlled trial making comparisons between PSIO clinical applications, or to no treatment, irrespective of language, date and status of publication. Every effort to decrease bias in the methodology employed was made. Screening, verification of eligibility, abstraction of information, assessment of risk of bias and of the quality of evidence were performed in duplicate, and any disagreement was resolved by discussion or consultation with the thesis co-supervisor until a final consensus was achieved. Finally, the random effects model was employed during exploratory quantitative data synthesis to incorporate any observed heterogeneity (Lau et al., 1997).

There are also some limitations to the present review, arising mainly from the nature and the characteristics of the data retrieved during the review process that resulted in the assessment of the level of available evidence as, at best, low. The scarcity of relevant high quality hierarchically evidence based information from RCTs, precluded meta-analytic procedures for most outcomes. Even in cases where such an approach was attempted, such quantitative syntheses can only be regarded as exploratory until additional research becomes available. However, current concepts support that data from even as few as two studies can be combined, provided that these can be meaningfully pooled (Ryan, 2013), as all other summarizing techniques are less transparent and/or are less likely to be valid (Valentine et al., 2010). Furthermore, exploratory subgroup analyses and analyses for “small-study effects” and publication bias (Higgins and Green, 2011), could not be carried out even though they were incorporated as possibilities according to the review protocol.

Another limitation of the data retrieved in this study stems from the small number of patients finally analyzed resulting an increased risk of bias and subsequent problems regarding the precision of the effect estimates in some cases. Moreover, it has to be acknowledged that the results of this review relates mostly to the comparison of the passive PSIO appliances to no treatment in Dutch infants with unilateral cleft lip and palate, since this constitutes the overwhelming majority of the information obtained. The inclusion criteria applied in the Dutchcleft study precludes the application of the results to other populations and procedures, thus diminishing the directness and generalizability of the available evidence. No studies on children with bilateral cleft lip and palate were retrieved. In addition, very limited data was locatable regarding active PSIO appliances or nasolaveloar molding techniques.

There was an additional limitation of the material included in this study originating from supplementary interventions carried out in some patients as part of the unavoidable individualization of treatment protocols required to cater for individual patient needs. However, the influence of these confounding parameters could not be clarified, either in the individual studies included in the present review, or as a part of subgroup analysis because of the lack of extensive relevant data. Finally, the involvement of different operators is also to be expected in long-term evaluations and can form a source of variability. This issue may constitute an additional reason of cautiousness when interpreting the results of this review.

6.4. Recommendations for future research

Evaluation of treatment modalities in the area of craniofacial abnormalities should be a continuous activity with use of proper scientific methodology (World Health Organization, 2002). As the overall quality of the relevant available evidence was considered at best as low, further research is imperative in order to discern more specifically the effects, if any, of various PSIO clinical applications, if present, and the relevant protocols on different parameters in the long term and to

be able to arrive at specific robust recommendations useful in the clinical setting. Well-designed and properly executed RCTs provide the best evidence with decreased risk of bias on the efficacy of health care interventions (Altman et al., 2001; Oxford Centre for Evidence-based Medicine, 2009). Long-term evaluation of outcomes of different treatment protocols is extremely valuable, since it is well known and understood that the definitive treatment outcomes in patients with cleft lip and/or palate cannot be recognized until facial development is complete (Niranjane et al., 2014).

To facilitate any type of research project thorough and proper records should be obtained from every patient, including traditionally orthodontics casts, facial photographs and radiographs. Recently, the use of three-dimensional records has been advocated. Also, a general agreement should be reached regarding the timing and type of records to be used. The EUROCRAN project (www.eurocran.org), advocates a fixed protocol regarding type and timing of records obtained. With a standardized set of records, collaboration and follow up can be simplified, while the inter-study comparison of outcomes may be facilitated by the use of meta-analytic techniques (Ellis, 2010).

Equally important, consensus has to be reached on the outcome measures relating to psychological variables and quality of life. Clinical trials are only as credible as their outcomes (Tugwell et al., 2007). A patient-centered approach, incorporating assessment of outcomes like patient satisfaction and the quality of life is a crucial part of treatment outcome and quality management and therefore it should be included in any relevant study (Sandy et al., 2012).

Furthermore, the vast majority of evidence-based data up to the present have involved a specific passive PSIO appliance protocol used in a particular Caucasian population with cleft lip and palate. Future research should also be directed to alternative PSIO protocols, as well as, populations of differing ethnic origin and patients with bilateral cleft lip and palate, before a better understanding of the different aspects of treatment can be achieved. In addition, every effort should be made to

increase the number of patients recruited, retained and analyzed in such studies, although the associated problems are well comprehended and expected. Increasing the number of patients analyzed will diminish the chance of a Type II error; at the same time it will increase the precision of effect estimates (Ellis, 2010). This will, in addition, facilitate the investigation of the effect of the confounding factors of individualized treatment interventions on the overall effect of PSIO appliances.

Finally, as resources are always limited in the context of health-care systems, investigation of the cost-effectiveness of the various treatment approaches is imperative (Severens et al., 1998) as is also the examination of possible adverse effects and problems related to PSIO appliances and procedures in order to enhance the quality of the delivered care (Prahl-Andersen et al., 2006).

7. CONCLUSIONS

Based on the findings of the present systematic review and meta-analysis, conducted following well-established guidelines, the null-hypotheses are not rejected. It seems that the investigated PSIO protocols used in patients with cleft lip and/or palate, generally, do not present significant effects when compared to each other or to no treatment, in terms of feeding characteristics and general body growth, facial esthetics, cephalometric variables, maxillary dentoalveolar variables and dental arch relationships, speech and language related variables, caregiver-reported outcomes, economic evaluation related outcomes, as well as, adverse effects and problems related to the appliances or the applied procedures. The aforementioned findings could provide initial guidance in the clinical setting, bearing in mind the small number of eligible trials, their heterogeneity with regards to treatment protocols, the results of the risk of bias assessment, as well as, the overall low quality of the available evidence. Given the multitude of parameters, which may have affected the results of the included trials, good practice would suggest further research in the respective field, in order to increase the body of high quality evidence and reach more robust relevant recommendations for management decisions in individual cases.

REFERENCES

Adams RA, Gordon C, Spangler AA. Maternal stress in caring for children with feeding disabilities: implications for health care providers. J Am Diet Assoc 1999;99:962-6.

Altalibi M, Humam S, Edwards R, Major PW, Flores-Mir F. Systematic review indices to assess malocclusions in patients with cleft lip and palate. Eur J Orthod 2013;36:772–82.

Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med 2001;134:663-94.

American Cleft Palate-Craniofacial Association. Parameters for evaluation and treatment of patients with cleft lip/palate or other craniofacial anomalies. Revised edition; 2009. Accessed at 31st of July 2016. Available at www.acpa-cpf.org.

Ardinger HH, Buetow KH, Bell GI, Bardach J, VanDemark DR, Murray JC. Association of genetic variation of the transforming growth factor-alpha gene with cleft lip and palate. Am J Hum Genet 1989;45:348-53.

Asher-McDade C, Roberts C, Shaw W C, Gallager C. Development of a method for rating nasolabial appearance in patients with clefts of the lip and palate. Cleft Palate-Craniofacial Journal 1991;28:385–91.

Athanasiou AE, Mazaheri M, Zarrinnia K. Frequency of crossbite in surgically treated cleft lip and/or palate children. J Pedod 1986;10:340-51.

Athanasiou AE, Mazaheri M, Zarrinnia K. Longitudinal study of dental arch dimensions in bilateral cleft lip and palate patients. J Pedod 1987a;11:253-68.

Athanasiou AE, Mazaheri M, Zarrinnia K. Longitudinal study of dental arch dimensions in hard and soft palate clefts. J Pedod 1987b;12:35-47.

Athanasiou AE, Mazaheri M, Zarrinnia K. Study of dental arch dimensions in unilateral cleft lip and palate patients. Cleft Palate J 1988;25:139-45.

Athanasiou AE, Moyers RE, Mazaheri M, Toutountzakis N. Frontal cephalometric evaluation of transverse dentofacial morphology and growth in children with isolated cleft palate. J Craniomaxillofacial Surg 1991b;19:249-53.

Athanasiou AE, Radhakrishnan S, Mazaheri M, Zarrinnia K. Transverse midface skeletal growth

in unilateral clefts of lip, alveolus and palate. In: Pfeifer G, ed. Craniofacial anomalies and clefts of lip, alveolus, and palate: Principles of treatment, long-term results. Stuttgart: Georg Thieme Verlag; 1991a:262-7.

Athanasiou AE, Tseng C-Y, Zarrinnia K, Mazaheri M. Frontal cephalometric study of transverse dentofacial growth in children with bilateral cleft of lip, alveolus and palate. J Craniomaxillofac Surg 1990;18:49-54.

Atherton JD. A descriptive anatomy of the face in human fetuses with unilateral cleft lip and palate. Cleft Palate J 1967;4:104–14.

Avedian LV, Ruberg RL. Impaired weight gain in cleft palate infants. Cleft Palate J 1980;17:24–6.

Bardach J, Morris HL, Olin WH. Late results of primary veloplasty: the Marburg project. Plast Reconstr Surg 1984;73:207-15.

Barillas I, Dec W, Warren ST, Cutting CB, Grayson BH. Nasoalveolar molding improves long-term nasal symmetry in complete unilateral cleft lip-cleft palate patients. Plast Reconstr Surg 2009;123:1002–6.

Baumrind S, Frantz RC. The reliability of head film measurements 1. Landmark identification. Am J Orthod 1971;60:111–27.

Becker M, Svensson H, Källén B. Birth weight, body length, and cranial circumference in newborns with cleft lip and palate. Cleft Palate Craniofaci J 1998;35:255–61.

Berkowitz S, Mejia M. A Comparison of the Effects of the Latham–Millard POPLA Procedure with a Conservative Treatment Approach on Dental Occlusion and Facial Aesthetics in CUCLP and CBCLP. In: Berkowitz S. Cleft Lip and Palate: Diagnosis and Management. Berlin: Springer; 2013:473-508.

Berkowitz S. Cleft Lip and Palate: Diagnosis and Management. Berlin: Springer; 2013a.

Berkowitz S. Cleft Lip and Palate. Berlin: Springer; 2013:453-71.

Berkowitz S. Facial and Palatal Growth. In: Berkowitz S. Cleft Lip and Palate: Diagnosis and Management. Berlin: Springer; 2013c:45-58.

Berkowitz S. The Cleft Palate Story. Berlin: Quintessence, Carol Stream; 1994:37-9.

Berkowitz S. The effect of clefting of the lip and palate and the palatal arch form. In: Berkowitz S. Cleft Lip and Palate: Diagnosis and Management. Berlin: Springer; 2013b:61-85.

Bessell A, Hooper L, Shaw WC, Reilly S, Reid J, Glenny AM. Feeding interventions for growth and development in infants with cleft lip, cleft palate or cleft lip and palate. Cochrane Database Syst Rev. 2011;2:CD003315.

Bongaarts CA, Kuijpers-Jagtman AM, van't Hof MA, Prahl-Andersen B. The effect of infant orthopedics on the occlusion of the deciduous dentition in children with complete unilateral cleft lip and palate (Dutchcleft). Cleft Palate Craniofac J 2004;41:633–41.

Bongaarts CA, Prahl-Andersen B, Bronkhorst EM, Prahl C, Ongkosuwito EM, Borstlap WA. Infant orthopedics and facial growth in complete unilateral cleft lip and palate until six years of age (Dutchcleft). Cleft Palate Craniofac J 2009;46:654–63.

Bongaarts CA, Prahl-Andersen B, Bronkhorst EM, Spauwen PHM, Mulder JW, Vaandrager JM, and Kuijpers-Jagtman AM. Effect of infant orthopedics on facial appearance of toddlers with complete unilateral cleft lip and palate (DUTCHCLEFT). Cleft Palate Craniofac J 2008;45:407–13.

Bongaarts CA, van't Hof MA, Prahl-Andersen B, Dirks IV, Kuijpers-Jagtman AM. Infant orthopedics has no effect on maxillary arch dimensions in the deciduous dentition of children with complete unilateral cleft lip and palate (Dutchcleft). Cleft Palate Craniofac J 2006;43:665–72.

Bongaarts CAM, van't Hof MA, Prahl-Andersen B, Kuijpers-Jagtman AM. Identification of cephalometric landmarks in unilateral cleft lip and palate patients: are there alternatives for point A, ANS and PNS? Cleft Palate Craniofac J 2008b;45:81–6.

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Metaanalysis. Chichester: Wiley 2009.

Boyles AL, DeRoo LA, Lie RT, Taylor JA, Jugessur A, Murray JC. Maternal alcohol consumption, alcohol metabolism genes, and the risk of oral clefts: a population- based case-control study in Norway, 1996–2001. Am J Epidemiol 2010;172:924–31.

Brophy TW. Cleft lip and cleft palate. J Am Dent Assoc 1927;14:1108–15.

Burdi AR, Silvey RG: Sexual differences in closure of the human palatal shelves. Cleft Palate J 1969;6:1–7.

Cash AC. Orthodontic Treatment in the Management of Cleft Lip and Palate. Front Oral Biol 2012;16:111–23.

Chan TC, Hayes C, Shusterman S, Mulliken JB, Will LA. The effects of active infant orthopedics on occlusal relationships in unilateral complete cleft lip and palate. Cleft Palate Craniofac J 2003;40:511–17.

Chang CS, Wallace CG, Pai BCJ, Chiu YT, Hsieh YJ, Chen IJ, Liao YF, Liou EJW, Chen PKT. Plastic and Reconstructive Surgery J 2014;134:275-82.

Cobourne MT. Cleft Lip and Palate. Epidemiology, Aetiology and Treatment. Front Oral Biol 2012.

Cobourne MT. The complex genetics of cleft lip and palate. Eur J Orthod 2004;26:7–16.

De la Vega A, Martinez E. Seasonal variation in the incidence of cleft lip and palate based on the age of conception. P R Health Sci J 2006;25:343–6.

Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M, Davey Smith G, Altman DG, eds. Systematic Reviews in Health Care. London: BMJ Books; 2001:285–312.

Der Simonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–88.

DiBiase D, Hunter S. A method of presurgical oral orthopaedics. Br J Orthod 1983;10:25–31.

Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet 2011;12:167–78.

Duncan PA, Shapiro LR, Soley RL, Turet SE. Linear growth patterns in patients with cleft lip or palate or both. Am J Dis Child 1983;137:159–63.

Ellis, Paul D. The Essential Guide to Effect Sizes: An Introduction to Statistical Power, Meta-Analysis and the Interpretation of Research Results. United Kingdom: Cambridge University Press. 2010.

Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behav Res Methods 2009;41:1149-60.

Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007;39:175-91.

Georgiade NG, Miadick RA, Thorne FL. Positioning of the premaxilla in bilateral cleft lips by oral pinning and traction. Plast Reconstr Surg 1968;41:240–43.

Georgiade NG. Early utilization of prosthetic appliances in cleft palate patients. Plast Reconstr Surg 1964;34:617–23.

Gnoinski WM. Infant orthopedics and later orthodontic monitoring for unilateral cleft lip and palate patients in Zurich. In: Bardach J, Morri s HL, eds. Multidisciplinary management of cleft lip and palate. Philadelphia: WB Saunders; 1990:578-85.

Gomez DF, Donohue ST, Figueroa AA, Polley JW. Nasal changes after presurgical nasoalveolar molding (PNAM) in the unilateral cleft lip nose. Cleft Palate Craniofac J. 2012;49:689–700.

Gorlin RJ, Cohen MM, Hennekam RCM, eds. Syndromes of the head and neck. New York: Oxford University Press; 2001.

Gray D, Mossey P. Evaluation of a modified Huddart/Bodenham scoring system for assessment of maxillary arch constriction in unilateral cleft lip and palate subjects. Eur J Orthod 2005;27:507–11.

Grayson BH, Cutting CB. Presurgical nasoalveolar orthopedic molding in primary correction of the nose, lip, and alveolus of infants born with unilateral and bilateral clefts. Cleft Palate Craniofac J 2001;38:193–8.

Grayson BH, Garfinkle JS. Nasoalveolar Molding and Columella Elongation in Preperation for the Primary Repair of Unilateral and Bilateral Cleft Lip and Palate. In: Losee JE, Kirschner RA, eds. Comprehensive Cleft Care. New York: McGraw Hill Medical; 2008:701-20.

Grayson BH, Maull D. Nasoalveolar molding for infants born with clefts of the lip, alveolus, and palate. Clin Plast Surg 2004;31:149–58.

Grayson BH, Santiago PE, Brecht LE, Cutting CB. Presurgical nasoalveolar molding in infants with cleft lip and palate. Cleft Palate Craniofac J 1999;36:486–98.

Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol 2011;64:380–2.

Hagberg C, Larson O, Milerad J. Incidence of cleft lip and palate and risks of additional malformations. Cleft Palate Craniofac J 1997;35:40–5.

Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*, version 5.1.0. The Cochrane Collaboration. 2011.

Hildebrandt KA, Fitzgerald HE. Adults' responses to infants varying in perceived cuteness. Behav Processes 1978;3:159–72.

Holland PW. Cloning and evolutionary analysis of msh-like homebox genes from mouse, zebrafish and ascidian. Gene 1991;98:253-7.

Hotz M, Gnoinski W, Perko M. The Zürich approach, 1964–1984. In: Hotz M, Gnoinski W, Perko M, eds. Early treatment of cleft lip and palate. Toronto: Huber; 1986:42–8.

Hotz M. Pre- and early post-operative growth-guidance in cleft lip and palate cases by maxillary orthopedics. Cleft Palate J 1969;6:361–72.

Hotz MM, Gnoinski WM. Comprehensive care of cleft lip and palate children at Zurich University: a primary report. Am J Orthod 1976;70:481-504.

Houston WJB. The analysis of errors in orthodontic measurements. Am J Orthod 1983;83:382–90.

Huddart AG. Treatment procedures in cleft lip and palate cases. Br Dent J 1967;122:185–92.

Ioannidis JP. Interpretation of tests of heterogeneity and bias in meta analysis. J Eval Clin Pract 2008;14:951-7.

Jensen BL, Dahl E, Kreiborg S. Longitudinal study of body height, radius length and skeletal maturity in Danish boys with cleft lip and palate. Scand J Dent Res 1983;91:473–81.

Jocelyn LJ, Penko MA, Rode HL. Cognition, communication and hearing in young children with cleft lip and palate and in control children: a longitudinal study. Pediatrics 1996;97:529–34.

Karling J, Larson O, Leanderson R, Henningsson G. Speech in unilateral and bilateral cleft palate patients from Stockholm. Cleft Palate Craniofac J 1993;30:73–7.

Konst EM, Prahl C, Weersink-Braks H, De Boo T, Prahl-Andersen B, Kuijpers- Jagtman AM. Cost-effectiveness of infant orthopedic treatment regarding speech in patients with complete unilateral cleft lip and palate: a randomized three center trial in the Netherlands (Dutchcleft). Cleft Palate Craniofac J 2004;41:71–7.

Konst EM, Rietveld T, Peters HF, Kuijpers-Jagtman AM. Language skills of young children with unilateral cleft lip and palate following infant orthopedics: a randomized clinical trial. Cleft Palate Craniofac J 2003b;40:356–62.

Konst EM, Rietveld T, Peters HF, Prahl-Andersen B. Phonological development of toddlers with unilateral cleft lip and palate who were treated with and without infant orthopedics: a randomized clinical trial. Cleft Palate Craniofac J 2003a;40:32–9.

Konst EM, Rietveld T, Peters HF, Weersink-Braks H. Use of a perceptual evaluation instrument to assess the effects of infant orthopedics on the speech of toddlers with cleft lip and palate. Cleft Palate Craniofac J 2003c;40:597–605.

Konst EM, Weersink-Braks H, Rietveld T, Peters H. An intelligibility assessment of toddlers with cleft lip and palate who received and did not receive presurgical infant orthopedic treatment. J Commun Disord 2000;33:483–501.

Konst EM, Weersink-Braks H, Rietveld T, Peters H. Prelexical development of unilateral cleft lip and palate babies with reference to presurgical infant orthopaedics: a randomized prospective clinical trial. Clin Linguist Phon 1999;13:395–407.

Konst EM. The Effects of Infant Orthopaedics on Speech and Language Development in Children with Unilateral Cleft Lip and Palate [Thesis]. Nijmegen: University of Nijmegen; 2002.

Kramer GJC, Hoeksma JB, Prahl-Andersen B. Early changes in complete cleft lip and/or palate. Acta Anat 1992;44:202-12.

Kuijpers-Jagtman AM, Prahl C. Neonatal Maxillary Orthopedics: Past to Present. In: Berkowitz S. Cleft Lip and Palate: Diagnosis and Management. Berlin: Springer; 2013:455-471.

Kuijpers-Jagtman AM, Ross EL. The influence of surgery and orthopedic treatment on maxillofacial growth and maxillary arch development in patients treated for orofacial clefts. Cleft Palate Craniofac J 2000;37:527–39.

Latham RA. Orthopedic advancement of the cleft maxillary segment: a preliminary report. Cleft Palate J 1980;17:227–33.

Lau J, Ioannidis JP, Schmid CH. Quantitative synthesis in systematic reviews. Ann Intern Med 1997;127:820–6.

Lee J, Nunn J, Wright C. Height and weight achievement in cleft lip and palate. Arch Dis Child 1997;76:70–2.

Liao YF, Hsieh YJ, Chen IJ, Ko WC, Chen PK. Comparative outcomes of two nasoalveolar molding techniques for unilateral cleft nose deformity. Plast Reconstr Surg 2012;130:1289–95.

Liou EJ, Subramanian M, Chen PK, Huang CS. The progressive changes of nasal symmetry and growth after nasoalveolar molding: A three-year follow-up study. Plast Reconstr Surg 2004;114:858–64.

Little J. Congenital anomalies and smoking. United Kingdom Department of Health. 2003.

Lohmander A, Lillvik M, Friede H. The impact of early infant jaw-orthopaedics on early speech production in toddlers with unilateral cleft lip and palate. Clin Linguist Phon 2004;18:259–84.

Long RE, Hathaway R, Daskalogiannakis J, Mercado A, Russell K, Cohen M. The americleft study: an inter-center study of treatment outcomes for patients with unilateral cleft lip and palate. part 1. Principles and study design. Cleft Palate Craniofac J 2011;48:239–43.

Marci V, Athanasiou AE. Source of error in lateral cephalometry. In: Athanasiou AE. Orthodontic Cephalometry, London: Mosby- Wolfe; 1995:125–40.

Masarei AG, Wade A, Mars M, Sommerlad BC, Sell D. A randomized control trial investigating the effect of presurgical orthopedics on feeding in infants with cleft lip and ⁄ or palate. Cleft Palate Craniofac J 2007;44:182–93.

Matsuo K, Hirose T, Otagiri T, Norose N. Repair of cleft lip with nonsurgical correction of nasal deformity in the early neonatal period. Plast Reconstr Surg 1989;83:25–31.

Maull DJ, Grayson BH, Cutting CB, Brecht LL, Bookstein FL, Khorrambadi D. Long-term effects of nasoalveolar molding on three-dimensional nasal shape in unilateral clefts. Cleft Palate Craniofac J 1999; 36:391–7.

Mazaheri M, Athanasiou AE, Long Jr RE, Kolokitha O. Evaluation of maxillary dental arch form in unilateral clefts of lip, alveolus and palate from one month to 4 years. Cleft Palate Craniofac J 1993;30:90-3.

McCance AM, Robert-Harry D, Sherriff M, Mars M, Houston WJB. A study model analysis of adult unoperated Sri Lankans with unilateral cleft lip and palate. Cleft Palate J 1990;27:146-54.

McNeil CK. Oral and Facial Deformity. London: Pitman and Sons; 1954:14-25.

McNeil CK. Orthodontic procedures in the treatment of congenital cleft palate. Dent Rec 1950;70:126–32.

Meyer-Gutknecht H, Fritz U, Schott TC. Methods to evaluate compliance of patients with removable appliances – survey results. J Orofac Orthop 2014;75:144-53.

Millard DR, Latham R, Huifen X, Spiro S, Morovic C. Cleft lip and palate treated by presurgical orthopedics, gingivoperiosteoplasty, and lip adhesion (POPLA) compared with previous lip adhesion method: a preliminary study of serial dental casts. Plast Reconstr Surg 1999;103:1630-44.

Millard DR. Cleft Craft. The Evolution of Its Surgery. The Unilateral Deformity. Philadelphia: Lippincott Williams and Wilkins; 1976.

Mishima K, Mori Y, Sugahara T, Minami K, Sakuda M. Comparison between palatal configurations in UCLP infants with and without a Hotz plate until four years of age. Cleft Palate Craniofac J 2000;37:185– 90.

Mishima K, Mori Y, Sugahara T, Sakuda M. Comparison between the palatal configurations in complete and incomplete unilateral cleft lip and palate infants under 18 months of age. Cleft Palate Craniofac J 2001;38:49–54.

Mishima K, Sugahara T, Mori Y, Sakuda M. Three-dimensional comparison between the palatal forms in infants with complete unilateral cleft lip, alveolus, and palate (UCLP) with and without Hotz's plate. Cleft Palate Craniofac J 1996a;33:245–51.

Mishima K, Sugahara T, Mori Y, Sakuda M. Three-dimensional comparison between the palatal forms in complete unilateral cleft lip and palate with and without Hotz plate from cheiloplasty to palatoplasty. Cleft Palate Craniofac J 1996b;33:312–7.

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. Open Medicine 2009;3:123–30.

Mølsted K, Asher-McDade C, Brattström V, Dahl E, Mars M, McWilliam J, Plint DA, Prahl-Andersen B, Semb G, Shaw WC, and The RPS. A six-center international study of treatment outcome in patients with clefts of the lip and palate: Part 2. Craniofacial form and soft tissue profile. Cleft Palate Craniofac J 1992;29:398–404.

Moore CC, MacDonald I, Latham R, Brandt MG. Septopalatal protraction for correction of nasal septal deformity in cleft palate infants. Otolaryngol Head Neck Surg 2005;133:949-53.

Mossey P, Clark JD, Gray D. Preliminary investigation of a modified Huddart/Bodenham scoring system for assessment of maxillary arch constriction in unilateral cleft lip and palate subjects. Eur J Orthod 2003;25:251–7.

Mossey P, Little J. Addressing the challenges of cleft lip and palate research in India. Indian J Plast Surg 2009;42:9–18.

Mossey PA, Little J, Munger RG, Dixon MJ, Shaw WC. Cleft lip and palate. Lancet 2009;374:1773–85.

Mossey PA, Little J. Epidemiology of oral clefts: an international perspective. In: Wyszynski DF, ed. Cleft Lip and Palate: From Origin to Treatment. New York: Oxford University Press; 2002;12:127–58.

Mossey PA, Modell B. Epidemiology of oral clefts 2012: an international perspective. Front Oral Biol 2012;16:1-18.

Murray JC. Gene/environment causes of cleft lip and/or palate. Clin Genet 2002; 61:248–56.

Murthy J. Management of cleft lip and palate in adults. Indian J Plast Surg 2009;42:116–22.

Neiman GS, Savage HE. Development of infants and toddlers with clefts from birth to three years of age. Cleft Palate Craniofac J 1997;3:218–5.

Niranjane PP, Kamble RH, Diagavane SP, Shrivastav SS, Batra P, Vasudevan SD. Current status of presurgical infant orthopaedic treatment for cleft lip and palate patients: A critical review. Indian J Plast Surg 2014;47:293-302.

Nollet PJPM, Katsaros C, Huyskens RWF, Borstlap WA, Bronkhorst EM, Kuijpers-Jagtman AM. Cephalometric evaluation of long-term craniofacial development in unilateral cleft lip and palate patients treated with delayed hard palatal closure. Int J Oral Maxillofac Surg 2008;37:123-30.

Noveraaz RLM, Disse MA, Ongkosuwito EM, Kuijpers-Jagtman AM. Transverse dental arch relationship at 9 and 12 years in children with unilateral cleft lip and palate treated with infant orthopedics: a randomized clinical trial (Dutchcleft). Clin Oral Invest 2015;19:2255-65.

Online Mendelian Inheritance in Man, OMIM®. Baltimore, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University; 2000. Accessed at 31st of July 2016. Available at http://omim.org.

Oxford Centre for Evidence-based Medicine – Levels of Evidence; 2009. Accessed at 31st of July 2016. Available at http://www.cebm.net/index.aspx?o=1025.

Panamonta V, Pradubwong S, Panamonta M, Chowchuen B. Global birth prevalence of orofacial clefts: A Systematic Review. J Med Assoc Thai 2015;98:11-21.

Papadopoulos MA, Koumpridou EN, Vakalis ML, Papageorgiou SN. Effectiveness of pre-surgical infant orthopedic treatment for cleft lip and palate patients: a systematic review and meta-analysis. Orthod Craniofac Res 2012;15:207–36.

Papageorgiou SN, Xavier GM, Cobourne MT. Basic study design influences the results of orthodontic clinical investigations. J Clin Epidemiol 2015;68:1512-22.

Paradise JL, Bluestone CD, Felder H. The universality of otitis media in 50 infants with cleft palate. Paediatrics 1969;44:35–42.

Prahl C, Kuijpers-Jagtman AM, van't Hof MA, Prahl-Andersen B. A randomized prospective clinical trial of the effect of infant orthopedics in unilateral cleft lip and palate: prevention of collapse of the alveolar segments (Dutchcleft). Cleft Palate Craniofac J 2003;40:337–42.

Prahl C, Kuijpers-Jagtman AM, van't Hof MA, Prahl-Andersen B. A randomised prospective clinical trial into the effect of infant orthopaedics on maxillary arch dimensions in unilateral cleft lip and palate (Dutchcleft). Eur J Oral Sci 2001;109:297–305.

Prahl C, Kuijpers-Jagtman AM, van't Hof MA, Prahl-Andersen B. Infant orthopedics in UCLP; effect on feeding, weight and length; a randomized clinical trial (Dutchcleft). Cleft Palate Craniofac J 2005;42:171-7.

Prahl C, Kuijpers-Jagtman AM, van't Hof MA, Prahl-Andersen B. Infant orthopedics and facial appearance: a randomized clinical trial (Dutchcleft). Cleft Palate Craniofac J 2006;43:659-64.

Prahl C, Prahl-Andersen B, van't Hof MA, Kuijpers-Jagtman AM. Presurgical orthopedics and satisfaction in motherhood: a randomized clinical trial (Dutchcleft). Cleft Palate Craniofac J 2008;45:284-8.

Prahl-Andersen B, Ju Q. Quality improvement of cleft lip and palate treatment. Angle Orthod 2006;76:265-8.

Proffit WR, Fields HW, Sarver DM. Contemporary Orthodontics, 4th edition, Orthodontic treatment planning: Limitations, controversies, and special problems. New Delhi: Mosby; 2007:287–92.

Pruzansky S. Pre-surgical orthopedics and bone grafting for infants with cleft lip and palate. A dissent. Cleft Palate J 1964;1:164-86.

Rahimov F, Jugessur A, Murray JC. Genetics of nonsyndromic orofacial clefts. Cleft Palate Craniofac J 2011;49:73–91.

Ranta R. A review of tooth formation in children with cleft lip/palate. Am J Orthod Dentofacial Orthop 1986;90:11–8.

Riski JE. Parents of children with cleft lip and plate. Clin Commun Disord 1991;1:42-7.

Ross RB, MacNamera MC. Effect of presurgical infant orthopedics on facial esthetics in complete bilateral cleft lip and palate. Cleft Palate Craniofac J 1994;31:68–73.

Ross RB. Treatment variables affecting facial growth in complete unilateral cleft lip and palate. Cleft Palate J 1987;24:5-77.

Ryan R. Cochrane Consumers and Communication Review Group. Cochrane Consumers and Communication Review Group: meta-analysis; 2013. Accessed at 31st of July 2016. Available at http://cccrg.cochrane.org.

Sandy J, Kilpatrick N, Ireland A. Treatment Outcome for Children Born with Cleft Lip and Palate. Front Oral Biol 2012;16:91–100.

Santiago PE, Grayson BH. Role of the craniofacial orthodontist on the craniofacial and cleft lip and palate team. Semin Orthod 2009;15:225–43.

Satokata I, Maas R. Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development. Nat Genet 1994;6:348-56.

Scott JH. The cartilage of the nasal septum. Br Dent J 1953;95:37–43.

Semb G. A study of facial growth in patients with uni- lateral cleft lip and palate treated by the Oslo CLP Team. Cleft Palate Craniofac J 1991;28:1-21.

Severens JL, Prahl C, Kuijpers-Jagtman AM, Prahl-Andersen B. Short-term cost-effectiveness analysis of presurgical orthopedic treatment in children with complete unilateral cleft lip and palate. Cleft Palate Craniofac J 1998;35:222–6.

Shaw WC, Asher-McDade C, Brattstro¨m V, Dahl E, McWilliam J, Mølsted K. A six-center international study of treatment outcome in patients with clefts of the lip and palate: part 1. Principles and study design. Cleft Palate Craniofac J 1992;29:393–7.

Shaw WC, Dahl E, Asher-McDade C, Brattstorm V, Mars M, McWilliam J, Molsted K, Plint DA, Prahl-Andersen B, Roberts CT, Semb G, The RPS. A six center international study of treatment outcome in patients with clefts of lip and palate: Part 5. General discussion and conclusions. Cleft Palate Craniofac J 1992;29:413-8.

Shaw WC, Semb G, Nelson P, Brattstorm V, Molsted K, Prahl-Andersen B. The Eurocleft project 1996-2000. Amsterdam: IOS Press; 2000.

Sierra FJ, Turner C. Maxillary orthopaedics in the presurgical management of infants with cleft lip and palate. Paediatric Dent 1995;17:419–23.

Solnit AJ, Stark MH. Mourning and the birth of a defective child. Psychoanal St Child 1961;16:9-24.

Sperber GH, Sperber SM. Embryogenetics of Cleft Lip and Palate. In: Berkowitz S, ed. Cleft Lip and Palate: Diagnosis and Management. Berlin: Springer; 2013:3-33.

Tobiasen JM, Hiebert JM, Boraz RA. Development of scales of severity of facial cleft impairment. Cleft Palate Craniofac J 1991;28:419–24.

Tobiasen JM. Scaling facial impairment. Cleft Palate J 1989;26:249–54.

Tobiasen JM. Social judgments of facial deformity. Cleft Palate J 1987;24:323–7.

Trpkova B, Major P, Prasad N, Nebbe B. Cephalometric landmark identification and reproducibility: a meta analysis. Am J Orthod 1997;112:165–70.

Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L. OMERACT: an international initiative to improve outcome measurement in rheumatology. Trials 2007;8:38.

Turner S R, Rumsey N, Sandy JR. Psychological aspects of cleft lip and palate. Euro J Orthod 1998;20:407–15.

Uzel A, Alparslan ZN. Long-term effects of presurgical infant orthopedics in patients with cleft lip and palate: a systematic review. Cleft Palate Craniofac J 2011;48:587-95.

Valentine JC, TD Pigott, HR Rothstein. How many studies do you need? A primer on statistical power for meta-analysis. J Educ Behav Stat 2010;35:215-47.

van den Boogaard MJ, Dorland M, Beemer FA, van Amstel HK. MSX1 mutation is associated with orofacial clefting and tooth agenesis in humans. Nat Genet 2000;24:342-3.

Vieira AR. Genetic and Environmental Factors in Human Cleft Lip and Palate. Front Oral Biol

2012;16:19–31.

Weatherley-White RC, Ben S, Jin Y, Riccardi S, Arnold TD, Spritz RA. Analysis of genomewide association signals for nonsyndromic cleft lip/palate in a Kenya African cohort. Am J Med Genet A 2011;155:2422–5.

Winters JC, Hurwitz DJ. Presurgical orthopaedics in the surgical management of unilateral cleft lip and palate. Plast Reconst Surg 1995;95:755–64.

Witzel MA, Salyer KE, Ross RB. Delayed hard palate closure: philosophy revisited. Cleft Palate J 1984;21:263-8.

World Health Organisation. Global strategies towards reducing the health-care burden of craniofacial anomalies. Report of WHO meetings on international collaborative research on craniofacial anomalies, WHO Human Genetics Programme. 2002.

Wu J, Zheng Q, Huang YQ, Wang Y, Li S, Lu DW. Significant evidence of association between polymorphisms in ZNF533, environmental factors, and nonsyndromic orofacial clefts in the Western Han Chinese population. DNA Cell Biol 2011;30:47–54.

Wu T, Liang KY, Hetmanski JB, Ruczinski I, Fallin MD, Ingersoll RG. Evidence of gene-environment interaction for the IRF6 gene and maternal multivitamin supplementation in controlling the risk of cleft lip with/without cleft palate. Hum Genet 2010;128:401–10.

Wyszynski DF, Duffy DL, Beaty TH. Maternal cigarette smoking and oral clefts: a meta-analysis. Cleft Palate Craniofac J 1997;34:206-10.

Zarrinnia K, Hulnick S, Athanasiou AE, Zaghi F. Taking dental impressions in infants with cleft lip and palate. Cleft Palate Craniofac J 1993;30:248.

Zucchero TM, Cooper ME, Maher BS. Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. N Engl J Med 2004;351:769-80.

Appendix I

Systematic review protocol used for registration with international prospective register of systematic reviews (PROSPERO).

Review question(s)

The aim of this study is to investigate the effectiveness of pre-surgical infant orthopedic (PSIO) appliances in patients with non-syndromic cleft lip and/or palate and evaluate the quality of the available evidence.

Searches

Comprehensive electronic database searches will be undertaken without language restriction in the following databases: MEDLINE via PubMed, the Cochrane Central Register of Controlled Trials (CENTRAL), Scopus, Web of Science, LILACS, IndMed, Scielo and Arab World Research Source. Unpublished literature will be accessed electronically using Google Scholar (https://scholar.google.com), ClinicalTrials.gov (http://clinicaltrials.gov), International Standard Randomised Controlled Trial Number (ISRCTN) registry (http://www.isrctn.com) and OpenGrey (http://www.opengrey.eu). In addition, Pro-Quest Dissertation and Theses Global database will be searched. Efforts will be made to obtain conference proceedings and abstracts where possible. Authors will be contacted to identify unpublished or ongoing clinical trials and to clarify methodology and data as necessary. Reference lists of included studies will be screened for additional relevant research.

Types of study to be included

The trials to be included should be RCTs evaluating outcomes of PSIO appliance treatment.

Condition or domain being studied

Patients with non-syndromic cleft lip and/or palate.

Participants/ population

Children of any age with any kind of non-syndromic cleft lip and/or palate defect.

Intervention(s), exposure(s)

Any type of PSIO appliance protocol.

Comparator(s)/ control

No treatment or alternative PSIO appliance protocol.

Outcome(s)

Primary outcomes

Feeding characteristics and nutritional status, facial esthetics, dentofacial cephalometric variables, maxillary dentoalveolar variables, dental arch relationships, hearing, speech and language evaluation.

Secondary outcomes

Patient and caregiver-reported outcomes, economic evaluation related outcomes, adverse effects and problems related to PSIO appliances and procedures.

Data extraction, (selection and coding)

All assessments including titles and/or abstract screening, full text evaluation, and extraction of data will be performed independently and in duplicate by two investigators (HRH and EGK). The

investigators will not be blinded to the authors or the results of the research. Disagreements will be resolved by discussion and consultation with a third author where necessary (AEA).

Risk of bias (quality) assessment

Assessment of risk of bias will be performed independently and in duplicate by two investigators (HRH and EGK) using the Cochrane Collaboration risk of bias tool that considers seven domains: random sequence generation; allocation concealment; blinding of participants and personnel; blinding of assessors; incomplete outcome data; selective reporting of outcomes; and other potential sources of bias. Each domain will receive a judgment of low, high or unclear risk of bias (indicating either lack of sufficient information to make a judgment or uncertainty over the risk of bias). Studies will be finally grouped into the following categories:

• low risk of bias (plausible bias unlikely to seriously alter the results): if all key domains of the study are at low risk of bias,

• unclear risk of bias (bias that raises some doubt about the results): if one or more key domains of the study are unclear, and,

• high risk of bias (bias that seriously weakens confidence in the results): if one or more key domains are at high risk of bias.

Disagreements will be resolved by discussion and consultation with a third author where necessary (AEA).

Strategy for data synthesis

Where studies have used the same type of intervention with the same outcome measure, we will pool the results using a random-effects meta-analysis analysis in view of the likely variation in population groups and settings. Depending on the variation in the indices used to quantify primary or secondary outcomes, we will use weighted or standardized mean differences for continuous

outcomes and risk ratios for binary outcomes, and calculate 95% confidence intervals and two sided p values for each outcome. Heterogeneity will be assessed using both the Chi-square test and the I-squared statistic.

Analysis of subgroups or subsets

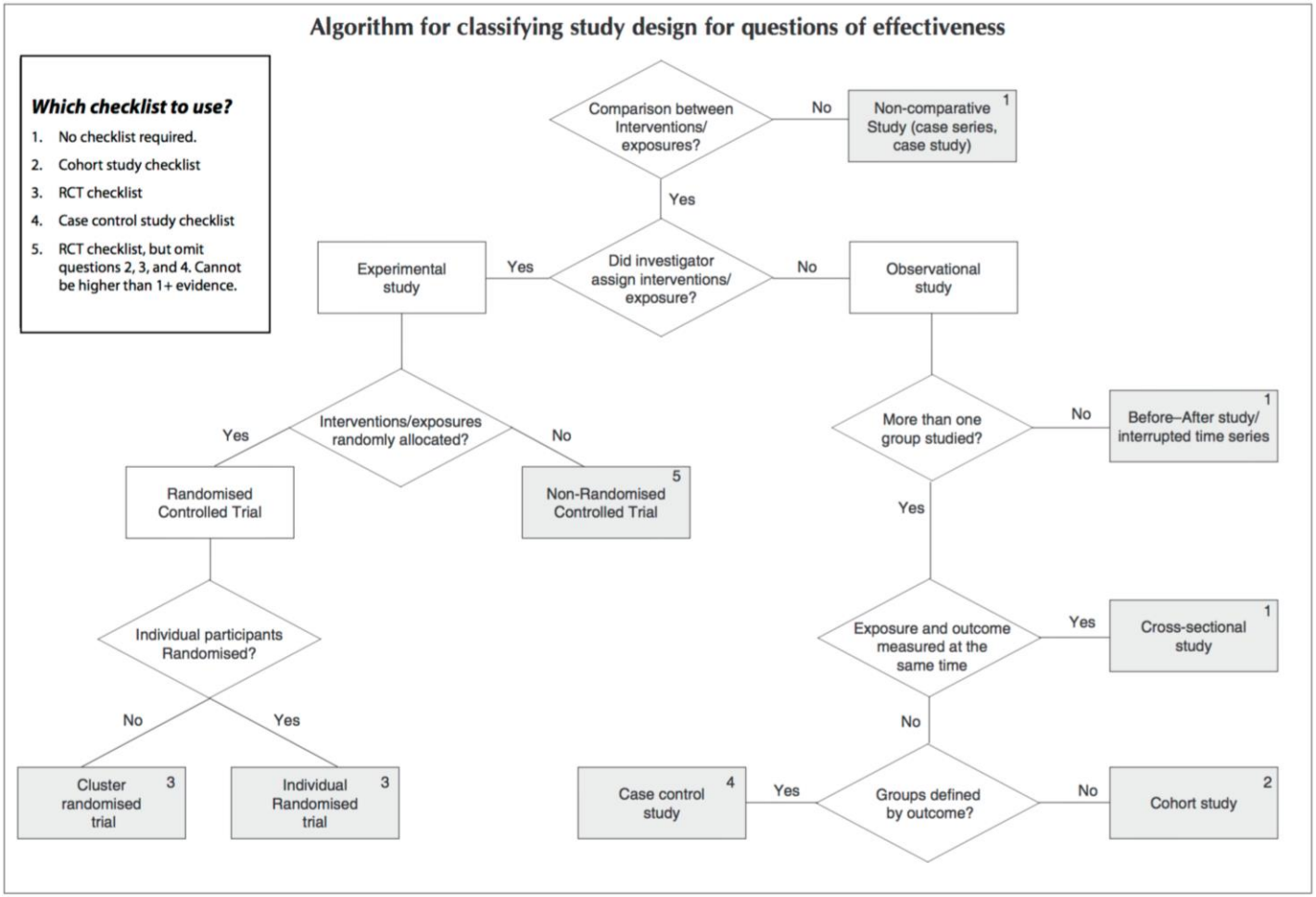
If possible, studies will be divided into categories depending on participants and intervention characteristics. If an adequate number of trials are identified, we will carry out analyses for "small-study effects" and publication bias.

Dissemination plans

Publication in peer reviewed journals.

Appendix II.

Scottish Intercollegiate Guidelines Network (SIGN) algorithm for classifying study design for questions of effectiveness



Appendix III

Strategy for database search [until May 1st, 2016].

Database	Search strategy	Hits
General Sources		
MEDLINE via PubMed http://www.ncbi.nlm.nih.gov/pubmed	(((randomized controlled trial[pt]) OR (controlled clinical trial[pt]) OR (randomized[tiab]) OR (placebo[tiab]) OR (drug therapy[sh]) OR (randomly[tiab]) OR (trial[tiab]) OR groups[tiab])) NOT ((animals[mh] NOT humans[mh])) AND ((cleft lip) OR (cleft-lip) OR (cleft lip and palate) OR (cleft-lip and palate) OR (alveolar cleft*)) AND (infant OR presurgical OR pre-surgical OR preoperative OR pre-operative OR early) AND (orthodon* OR orthopedic* OR orthopaedic* OR nasoalveolar OR moulding OR molding OR pnam OR nam OR plate*)	**173**
Cochrane Central Register of Controlled Trials http://onlinelibrary.wiley.com/cochranelibrary	((cleft lip) OR (cleft-lip) OR (cleft lip and palate) OR (cleft-lip and palate) OR (alveolar cleft*)) AND (infant OR presurgical OR pre-surgical OR preoperative OR pre-operative OR early) AND (orthodon* OR orthopedic* OR orthopaedic* OR nasoalveolar OR moulding OR molding OR pnam OR nam OR plate*) in Title, Abstract, Keywords in Trials	**46**
Scopus https://www.scopus.com/	((randomized controlled trial) OR (controlled clinical trial) OR randomized OR placebo OR drug therapy OR randomly OR trial OR groups) AND ((cleft lip) OR (cleft-lip) OR (cleft lip AND palate) OR (cleft-lip AND palate) OR (alveolar cleft*)) AND (infant OR presurgical OR pre-surgical OR preoperative OR pre-operative OR early) AND (orthodon* OR orthopedic* OR orthopaedic* OR nasoalveolar OR moulding OR molding OR pnam OR nam OR plate*) AND (LIMIT-TO (SUBJAREA, "DENT")) AND (LIMIT-TO (EXACTKEYWORD , "Human") OR LIMIT-TO (EXACTKEYWORD , "Humans"))	**345**
Web of Science™ Core Collection http://apps.webofknowledge.com/	**TOPIC:** (((randomized controlled trial) OR (controlled clinical trial) OR randomized OR placebo OR drug therapy OR randomly OR trial OR groups) AND ((cleft lip) OR (cleft-lip) OR (cleft lip AND palate) OR (cleft-lip AND palate) OR (alveolar cleft*)) AND (infant OR presurgical OR pre-surgical OR preoperative OR pre-operative OR early) AND (orthodon* OR orthopedic* OR orthopaedic* OR nasoalveolar OR moulding OR molding OR pnam OR nam OR plate*)) **Refined by: RESEARCH AREAS:** (DENTISTRY ORAL SURGERY MEDICINE) **Timespan:** All years. Search language=Auto	**140**
Regional sources		
LILACS http://lilacs.bvsalud.org/en/	tw:(orthodon*) AND (instance:"regional") AND (db:("LILACS") AND type_of_study:("clinical_trials") AND limit:("humans"))	**85**
IndMed http://indmed.nic.in/indmed.html	orthodontic AND cleft	**10**
Arab World Research Source [2016 05 01] http://0-search.ebscohost.com.amclb.iii.com	orthodont* AND cleft*	**6**
Grey literature sources		
Google Scholar https://scholar.google.com	allintitle: randomized cleft Excluding patents & citations	**36**
ClinicalTrials.gov http://clinicaltrials.gov/	cleft lip | Studies With Results | Interventional Studies	**2**
ISRCTN registry http://www.isrctn.com	((randomized controlled trial) OR (controlled clinical trial) OR randomized OR placebo OR drug therapy OR randomly OR trial OR groups) AND ((cleft lip) OR (cleft-lip) OR (cleft lip AND palate) OR (cleft-lip AND palate) OR (alveolar cleft) OR (alveolar clefts)) AND (infant OR presurgical OR pre-surgical OR preoperative OR pre-operative OR early) AND (orthodontic	**4**

Database	Search strategy	Hits
	OR orthodontics OR orthopedic OR orthopedics OR orthopaedic OR Orthopaedics OR nasoalveolar OR moulding OR molding OR pnam OR nam OR plate OR plates)	
OpenGrey http://www.opengrey.eu/	((randomized controlled trial) OR (controlled clinical trial) OR randomized OR placebo OR drug therapy OR randomly OR trial OR groups) AND ((cleft lip) OR (cleft-lip) OR (cleft lip AND palate) OR (cleft-lip AND palate) OR (alveolar cleft*)) AND (infant OR presurgical OR pre-surgical OR preoperative OR pre-operative OR early) AND (orthodon* OR orthopedic* OR orthopaedic* OR nasoalveolar OR moulding OR molding OR pnam OR nam OR plate*)	**0**
ProQuest Dissertations and Theses Global http://search.proquest.com	ti((cleft lip) OR (cleft-lip) OR (cleft lip and palate) OR (cleft-lip and palate) OR (alveolar cleft*))	**184**

Appendix IV

Details of risk of bias assessment – Publications from the DUTCHCLEFT [Domains examined: 1: Random sequence generation 2: Allocation concealment, 3: Blinding of participants and personnel, 4: Blinding of outcome assessment, 5: Incomplete outcome data, 6: Selective outcome reporting, 7: Other potential threats to validity]

Study	Rating	Reasons for rating
DUTCHCELFT [General assessment]	**1.** Low	The exact method of randomization is mentioned. ["...subjects were randomly allocated to groups using the minimization methods..."]
	2. Low	Minimization is a convincing method of allocation concealment. ["...subjects were randomly allocated to groups using the minimization methods..."]
	3. Low	Blinding was not possible. However, the review authors believe that the outcome is not likely to be influenced by lack of blinding.
Bongaarts et al., 2004	**4.** Low	Statement that the investigator was blinded. No other reason to infer that the investigator could assume group allocation. ["To eliminate bias, all models were duplicated and trimmed in the same way. In this way the examiners were not able to identify a patient o r a cleft palate center."]
	5. Unclear	Dropouts are described and explained. However, it is unclear how they could influence the various outcomes of the study.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear whether compliance or the other performed interventions until the age of 6 years old introduced some kind of bias.
Bongaarts et al., 2006	**4.** Low	No reason to infer that the investigator could assume group allocation. ["To eliminate bias, all models were duplicated and trimmed in the same way. In this way the examiners were not able to identify a patient or a cleft palate center."]
	5. Unclear	Dropouts are described and explained. However, it is unclear how they could influence the various outcomes of the study.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear whether compliance or the other performed interventions until the age of 6 years old introduced some kind of bias.
Bongaarts et al., 2008	**4.** Low	No reason to infer that the investigators could assume group allocation.
	5. Unclear	Dropouts are described and explained. However, it is unclear how they could influence the various outcomes of the study.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear whether compliance or the other performed interventions until the age of 6 years old introduced some kind of bias.
Bongaarts et al., 2009	**4.** Low	No reason to infer that the investigators could assume group allocation.
	5. Unclear	Dropouts are described and explained. However, it is unclear how they could influence the various outcomes of the study.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear whether compliance, other performed interventions until the age of 6 years old, or uncertainly in landmark localization introduced some kind of bias.

Study	Rating	Reasons for rating
Konst et al., 1999	**4.** Low	No reason to infer that the investigators could assume group allocation.
	5. High	Dropouts are not fully described and explained. Moreover, it is unclear how they could influence the various outcomes of the study.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear how compliance may have influenced the results of the study.
Konst et al., 2000	**4.** Low	No reason to infer that the investigators could assume group allocation.
	5. High	Dropouts are not fully described and explained. Moreover, it is unclear how they could influence the various outcomes of the study.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear how compliance may have influenced the results of the study.
Konst et al., 2002 [only for future speech therapy need]	**4.** Low	No reason to infer that the investigators could assume group allocation.
	5. High	Dropouts are not fully described and explained. Moreover, it is unclear how they could influence the various outcomes of the study.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear how compliance may have influenced the results of the study.
Konst et al., 2003a	**4.** Low	No reason to infer that the investigators could assume group allocation.
	5. High	Dropouts are not fully described and explained. Moreover, it is unclear how they could influence the various outcomes of the study.
	6. Unclear	An outcome included in the original thesis publication has been omitted from the article.
	7. Unclear	It is unclear how compliance may have influenced the results of the study.
Konst et al., 2003b	**4.** Low	No reason to infer that the investigators could assume group allocation.
	5. High	Dropouts are not fully described and explained. Moreover, it is unclear how they could influence the various outcomes of the study.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear whether compliance or the other performed interventions until the age of 6 years old introduced some kind of bias.
Konst et al., 2003c	**4.** Low	No reason to infer that the investigators could assume group allocation.
	5. High	Dropouts are not fully described and explained. Moreover, it is unclear how they could influence the various outcomes of the study.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear how compliance may have influenced the results of the study.

Study	Rating	Reasons for rating
Konst et al., 2004	**4.** Low	No reason to infer that the investigators could assume group allocation.
	5. High	Dropouts are not fully described and explained. Moreover, it is unclear how they could influence the various outcomes of the study.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear how compliance may have influenced the results of the study.
Noverraz et al., 2015	**4.** Low	Statement that the investigator was blinded. No other reason to infer that the investigator could assume group allocation. ["To eliminate bias, all models were duplicated and trimmed in the same way. In this way the examiners were not able to identify a patient or a cleft palate center."]
	5. Unclear	Dropouts are described and explained. However, it is unclear how they could influence the various outcomes of the study.
	6. Unclear	Outcomes such as overjet, overbite and sagittal occlusion appearing in previous similar study from the same sample are missing.
	7. Unclear	It is unclear whether compliance or the other performed interventions until the age of 12 years old introduced some kind of bias.
Prahl et al., 2001	**4.** Low	No reason to infer that the investigator could assume group allocation. ["Data entry and analysis were blinded. The orthodontists who treated the patients did not assess the maxillary dimensions."]
	5. Low	Dropouts are described and explained.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear how compliance may have influenced the results of the study.
Prahl et al., 2003	**4.** Low	No reason to infer that the investigator could assume group allocation.
	5. Low	Dropouts are described and explained.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear how compliance may have influenced the results of the study.
Prahl et al., 2005	**4.** Low	No reason to infer that the investigator could assume group allocation.
	5. Unclear	Dropouts are described and explained. However, it is unclear how they could influence the various outcomes of the study.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear how compliance may have influenced the results of the study.
Prahl et al., 2006	**4.** Low	No reason to infer that the investigator could assume group allocation.
	5. Low	Dropouts are described and explained.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear how compliance may have influenced the results of the study.

Study	Rating	Reasons for rating
Prahl et al., 2008	**4.** Low	No reason to infer that the investigator could assume group allocation.
	5. Unclear	Dropouts are described and explained. However, it is unclear how they could influence the various outcomes of the study.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	The study questionnaire had not been previously validated.
Severens et al., 1998	**4.** Low	No reason to infer that the investigators could assume group allocation.
	5. High	Dropouts are not fully described and explained. Moreover, it is unclear how they could influence the various outcomes of the study.
	6. Low	All important outcomes are adequately reported.
	7. Low	The study appears to be free of other sources of bias.

Appendix V

Details of risk of bias assessment – Remaining studies. [Domains examined: 1: Random sequence generation 2: Allocation concealment, 3: Blinding of participants and personnel, 4: Blinding of outcome assessment, 5: Incomplete outcome data, 6: Selective outcome reporting, 7: Other potential threats to validity]

Study	Rating	Reasons for rating
Chang et al., 2014	**1.** Low	The exact method of randomization is mentioned. ["They were block randomized at a 1:1 ratio to undergo either modified Figueroa or modified Grayson nasoalveolar molding by an independent third-party specialized trials nurse using secure randomization envelopes."
	2. Low	No statement on the envelopes being sequentially administered. No other reason to infer that the investigator could influence group allocation. ["by an independent third-party specialized trials nurse using secure randomization envelopes."]
	3. Low	Blinding of the participants, caregivers and personnel was not possible. However, the review authors believe that the outcome is not likely to be influenced by lack of blinding.
	4. Low	Statement that the investigator was blinded. No other reason to infer that the investigator could assume group allocation. ["All assessors were blinded regarding the nasoalveolar molding technique that patients had been randomized to undergo; no breaks in the blind were reported."]
	5. Low	No dropouts occurred.
	6. Low	All important outcomes are adequately reported.
	7. Low	The study appears to be free of other potential threats to validity.
Masarei et al. 2007	**1.** Low	The exact method of randomization is mentioned. ["Minimization was therefore used to ensure that the Iwo groups contained similar numbers of first-born, later born, and male and female infants. Data for patient allocation were entered by the researcher using MINIM."]
	2. Low	Minimization is a convincing method of allocation concealment. ["Minimization was therefore used to ensure that the Iwo groups contained similar numbers of first-born, later born, and male and female infants. Data for patient allocation were entered by the researcher using MINIM."]
	3. Low	Blinding of the participants, caregivers and personnel was not possible. However, the review authors believe that the outcome is not likely to be influenced by lack of blinding.
	4. Low	No statement, but no reason to infer that the investigator could assume group allocation.
	5. Unclear	Dropouts are described and explained. However, it is unclear how they could influence the various outcomes of the study.
	6. Low	All important outcomes are adequately reported.
	7. Unclear	It is unclear how compliance may have influenced the results of the study.