The TOPS trial

Timing of Primary Surgery for Cleft Palate (TOPS): a randomised trial of palate surgery at age 6 months versus age 12 months

ClinicalTrials.gov Identifier: NCT00993551
MREC Number: 09/H1306/119
NIDCR protocol number: 09-017-E
Funder: National Institutes of Health (5U01DE018664/1U01DE018837)
Sponsor: The University of Manchester (57201)

Statistical Analysis Plan v3.0 12/03/2021

	ORIGINATED BY	QC PERFORMED BY	APPROVED BY
Name	Rachael Cooper	N/A	Paula Williamson
Title	Blinded Trial Statistician	N/A	Independent Statistician
Date	12/03/2021		
Protocol Version and Date	Version 6.0 (31/03/2020) Superseded versions: V1.0 (02/11/2009), V2.0 (10/03/2010) V3.0 (01/05/2013), V4.0 (26/08/2015), Version 5.0 (22/08/2018)		

1 Change control

Protocol version	Updated SAP version no.	Section number changed	Description of change	Date changed
5.0	2.0	Throughout	V1.0 of the SAP has required a major revision due to developing speech methodology, appropriate for the trial, that has emerged during trial conduct. V2.0 has therefore changed significantly throughout the document to reflect this new methodology and has been updated to be in line with the current Standard Operating Procedure template.	06/02/2020
6.0	3.0	Throughout	Minor clarification to wording has been added and typing errors corrected.	09/03/2021
6.0	3.0	14.1	Wording added to CONSORT diagram to summarise the level of withdrawal for those withdrawn from follow up. This was a recommendation from blind review	09/03/2021
6.0	3.0	17.2	Footnote added to Table 4 to clarify recordings taken as video, however assessments made on spontaneous speech or bus story recordings at 5 year where audio only due to remote assessment.	09/03/2021
6.0	3/0	17.3 and 18.6.2.9	Updated to allow right lateral photograph where provided and left is not available, no clinical reason to use left only. Addition to report number assessed on each side.	09/03/2021
6.0	3.0	18.5.2 and 23.5	Wording added to provide detail for the reassessments of words at 5 years following feedback from the initial assessment. Clarification on speech outcome derivations have been added following review of code ran on calibration data to ensure derivation is correct. Mapping structures at 3 and 5 year updated.	09/03/2021

Protocol version	Updated SAP version no.	Section number changed	Description of change	Date changed
6.0	3.0	18.1 and 18.6.1.3.3	Two sensitivity analysies have been added instead of a per protocol analysis. This was a recommendation from blind review. The sensitivity analysies will remove patients with the follwoing major protocol deviations from the primary outcome set: • Infants with syndromic cleft palate or severe developmental delay - Identification post randomisation • Surgery outside time window	09/03/2021
6.0	3.0	18.6.2.1.2	Wording added to explain that results will be presented as medians and IQRs, with a Mann Whitney U tests performed for secondary outcome 1a as data is skewed. This was a recommendation from blind review.	09/03/2021
6.0	3.0	18.6.2.2.2.1	Wording added to the definition of agreement to clarify the handling of 'could not assess' assessments for reliability.	09/03/2021
6.0	3.0	18.6.2.2.2	Wording added to explain that results will be presented as medians and IQRs, with a Mann Whitney U tests performed for secondary outcome 2b as data is skewed. This was a recommendation from blind review.	09/03/2021
6.0	3.0	18.6.2.3.2	Wording added to explain that results will be presented as means and SDs, with a t-test performed for secondary outcome 3b and 3c as data is normally distributed. This was a recommendation from blind review.	09/03/2021
6.0	3.0	18.6.2.4.1	Wording added to Table 4 to ensure outcomes reported as percentages.	09/03/2021
6.0	3.0	18.6.2.4.2	Wording added to explain that results will be presented as medians and IQRs, with a Mann Whitney U tests performed on each part (a to e) of secondary outcome 4 as data is skewed. This was a recommendation from blind review	09/03/2021
6.0	3.0	18.6.2.5.2	Wording added to explain that results will be presented as medians and IQRs, with a Mann Whitney U tests performed on each part (a to e) of secondary outcome 5 as data is skewed. This was a recommendation from blind review.	09/03/2021

Protocol version	Updated SAP version no.	Section number changed	Description of change	Date changed
6.0	3.0	18.6.2.7.1	Wording amendment to use visual reinforced audiometry where provided rather than sound field audiometry for secondary outcome 7 a ii at 12 months. This was a request from the lead audiologist following blind review.	09/03/2021
6.0	3.0	18.6.2.8	Wording added to allow classification according to the worst ear to summarise type A and C separately. The worst ear is to be classified as Type B, C then A. This was a recommendation from blind review.	09/03/2021
6.0	3.0	18.6.2.9.2	Wording added to explain that results will be presented as means and SDs, with a t-test performed for secondary outcome 9a and 9b as data is normally distributed. This was a recommendation from blind review.	09/03/2021
6.0	3.0	18.6.2.9.1	Wording added for the manual scoring of models where automated scoring could not be done due to software not picking up landmarks. Wording added to clarify reasons for failing quality assessments are reclassified by the dentofacial lead.	09/03/2021
6.0	3.0	18.6.2.10.2	Wording added to explain that results will be presented as means and SDs, with a t-test performed for secondary outcome 10a, 10b and 10c as data is normally distributed. This was a recommendation from blind review.	09/03/2021
6.0	3.0	20.3	Wording added to explain that results will be presented as medians and IQRs, with a Mann Whitney U tests performed for nasometry as data is skewed. This was a recommendation from blind review	09/03/2021
6.0	3.0	21.2.1	Following query responses some surgeons took the wrong measurement for 'bony width at posterior hard palate', these will be summarised as a cleft dimension separately to 'bony width at posterior hard palate'.	09/03/2021
6.0	3.0	21.2.4	Categories amended to include detail of treatment for respiratory in line with the other reported categories in Observations monitored 48 hours post surgery and data collected on CRF	09/03/2021

2 Approval and agreement

At a minimum two versions of the SAP should be approved and stored within the statistics trial file.

1. SAP version 1.0 should be created after it has been reviewed and signed-off to ensure all are in agreement with the planned analysis and no further changes are foreseen.
2. The final SAP version should be converted to PDF and signed following the blinded review for protocol deviations and immediately prior to database lock as evidence of the analysis planned prior to unbinding of the study.

SAP Version Number being approved: 3.0

Trial Statistician

Name _____

Signed _____ Date _____

Senior Statistician* [Senior statistician has seen unblinded data so has not written this SAP. Duty delegated to independent statistician who has not seen unblinded data

Name _____

Signed _____ Date _____

Chief Investigator/clinical lead

Name _____

Signed _____ Date _____

3 Roles and responsibilities

Rachael Cooper (Department of Health Data Science, University of Liverpool), Trial Statistician; Beth Conroy (Department of Health Data Science, University of Liverpool), Senior Trial Statistician; Carrol Gamble (Department of Health Data Science, University of Liverpool), Lead Statistician; P Williamson (Department of Health Data Science, University of Liverpool): Independent Statistician, William Shaw (School of Dentistry, University of Manchester) and Kevin Munro (University of Manchester), Chief Investigator(s); Tri Tat (Department of Biostatistics, University of Liverpool), Trial Statistician until April 2015.

Author's contributions

Tri Tat proposed and prepared the statistical analysis plan (V1.0). V1.0 required a major revision due to developing methodology, V2.0 of the analysis plan has been written blind to any analysis of comparative speech outcomes. Following the emergence of developing methodology, Rachael Cooper updated and prepared the statistical analysis plan (V2.0). Beth Conroy, Carrol Gamble and Kevin Munro (on behalf of William Shaw) read, amended and approved the statistical analysis plan (V2.0). Rachael Cooper updated the SAP in relation to recommendations made during blind review of the data. P Williamson read and reviewed the SAP and helped with questions relating to the analysis. Kevin Munro read, reviewed and approved the SAP.

4 List of abbreviations and definitions of terms

95% CI	95% confidence interval
AC	Administrative centre
AE	Adverse event
ANB	Angle between soft tissue nasion, A point, and B point on profile photograph
CB	Canonical babbling
CBR	Canonical babbling ratio
CInv	Chief investigator
CRF	Case report form
CSG	Core Speech Group
DSMB	Data Safety Monitoring Board
ITT	Intention to treat
IQR	Inter-quartile range
PCC	Percent consonants correct
PCM	Percent correct manner
PCP	Percent correct placement
PI	Principal Investigator
SAE	Serious adverse event
SD	Standard deviation
SLT	Speech and language therapist
TEOAE	Transient Otoacoustic Emission
TSC	Trial Steering Committee
UPR	Unanticipated problem
VPC-sum	Velopharyngeal composite score summary
VPC	Velopharyngeal function
VRA	Visual Reinforcement Audiometry

Contents

1 Change control 1
2 Approval and agreement 4
3 Roles and responsibilities 5
4 List of abbreviations and definitions of terms 6
5 Statement of compliance 10
6 Background and rationale 10
7 Study objectives 11
8 Investigational plan and study design 13
8.1 Description of overall study design and plan 13
8.2 Treatments studied 13
8.3 Treatment compliance 14
8.4 Patient population studied 14
8.4.1 Inclusion criteria 14
8.4.2 Exclusion criteria 15
8.4.3 Removal of patients from therapy or assessment 15
8.5 Consent process 16
8.6 Blinding 16
8.7 Method of assignment to treatment 16
8.8 Sequence and duration of all study periods 17
8.9 Schedule of assessments 17
9 Listing of outcomes 17
9.1 Primary outcome 17
9.2 Secondary outcomes 17
10 Determination of sample size 19
11 Study framework 19
12 Confidence intervals, p-values and multiplicity 20
13 Timing and objectives of interim and final analyses 20
13.1 Interim monitoring and analyses 20
13.2 Interim analysis at the end of the initial funding period 20
13.3 Final analysis 21
14 Disposition of participants 21
14.1 Screening, eligibility and recruitment 21
14.2 Post randomisation discontinuations 23
15 Protocol deviations 24

16 Unblinding 25
17 Completeness of follow up assessments 25
17.1 Case report forms 26
17.2 Speech recordings 28
17.3 Photographs 30
17.4 Impressions 30
18 Efficacy evaluations 31
18.1 Data sets analysed 31
18.2 Demographic and other baseline characteristics 32
18.3 Compliance with treatment 33
18.4 Compliance with study follow up and assessments 34
18.4.1 Post-operative assessments 35
18.4.2 12 month assessments 35
18.4.3 3 year assessments 37
18.4.4 5 year assessments 39
18.5 Software for speech outcome assessment 40
18.5.1 TimeStamper 40
18.5.2 SPEAK 43
18.6 Analysis of outcomes 52
18.6.1 Primary outcome: Velopharyngeal function at five years of age 52
18.6.2 Secondary outcomes 59
19 Missing data and withdrawals 96
20 Additional analyses 97
20.1 DENVER II 97
20.2 Speech therapy outside of routine trial visits 97
20.3 Nasometry 99
20.4 Secondary surgery 100
21 Safety evaluations 101
21.1 Data sets analysed 101
21.2 Presentation of the data 101
21.2.1 Baseline surgery characteristics 102
21.2.2 Intra operative events 102
21.2.3 Early complications during hospital stay 103
21.2.4 Observations monitored 48 hours post surgery 103
21.2.5 Post operative medication 104
21.2.6 Late complications 104

21.2.7 Serious adverse events 104
21.2.8 Unanticipated problems 106
22 References 107
23 Appendix 110
23.1 Definition of protocol deviations 110
23.2 Summary of outcomes 113
23.3 TimeStamper example output 120
23.4 SPEAK example output 121
23.4.1 3 year 121
23.4.2 5 year 123
23.5 Mapping structures 126

5 Statement of compliance

This Statistical Analysis Plan (SAP) provides a detailed and comprehensive description of the pre-planned **final** analyses for the study "The **TOPS** trial: **T**iming **O**f **P**rimary **S**urgery for cleft palate". The planned statistical analyses described within this document are compliant with those specified in the brief within the TOPS Protocol version 5.0 (22/08/2018). This SAP comprehensively describes in the planned final analysis.

This study is carried out in accordance with the World Medical Association Declaration of Helsinki (1964) and the Tokyo (1975), Venice (1983), Hong Kong (1989) and South Africa (1996) amendments and will be conducted in compliance with the TOPS Protocol, Liverpool Clinical Trials Centre (LCTC) Standard Operating Procedures (SOPs) and EU Directive 2001/20/EC, transposed into UK law as the UK Statutory Instrument 2004 No 1031: Medicines for Human Use (Clinical Trials) Regulations 2004.

These planned analyses will be performed primarily by the trial statistician, with assistance from other members of the statistics team where required.

The results of the final analysis described within this statistical analysis plan will be contained within a statistical analysis report. This report will be used as the basis of the primary research publications according to the study publication plan.

All analyses will be performed with standard statistical software (SAS 9.4 or later). The finalised analysis datasets, programs and outputs will be archived following Good Clinical Practice guidelines and standard operating procedure SOP GE012 Archiving Procedures in CTRC. The testing and validation of the statistical analysis programs will be performed following SOP ST001.

6 Background and rationale

The background and rationale for this trial is outlined in depth in the TOPS Protocol.

To summarise, clefts of the lip and/or palate, occurring with an incidence of about 1 per 600 births, are among the most common birth anomalies. This trial focuses on isolated clefts of the palate which occur with a global incidence of 4.5 per 10,000 births[1]. Depending on geographic location, the prevalence of isolated clefts of the palate ranges from 1.8 to 14.6 per 10,000 [1].

The timing of palatal surgery has been a controversial issue since the 1930s [2]. Traditionally, rationale for delaying hard palate surgery was partly based on the belief that postponing the trauma of palatal closure may reduce maxillary growth disturbance. However, there is little evidence that facial skeletal growth in individuals with isolated cleft palate is substantially affected by different surgical protocols, though maxillary arch form, especially transversely, may be affected [3-6].

Over recent decades, age at palatal surgery in infancy has fallen. This has led to one stage palatal closure around the age of 12 months, or as early as 6 months, becoming standard practice at cleft units in Europe and the USA. Protagonists of early closure of the palatal cleft have proposed that since speech is a learned behaviour, the sooner an intact anatomy is created, the better [7-10]. As yet however, there is no evidence that better speech is achieved.

The widespread uncertainty surrounding the timing of palate closure was reflected in the diversity of protocols currently employed by the partners in the Scandcleft Research Group. This Group, a partnership of Scandinavian and UK cleft lip and palate centres [11-22], considered this to be the priority question for a future trial. Its aim was to determine whether, in infants with cleft palate, repair at either age 6 or at 12 months (corrected for gestational age) would achieve better speech outcomes. The design of the trial was supported by a planning grant from the National Institutes of Health, who subsequently funded the proposed trial.

7 Study objectives

The primary research objective is to determine whether surgery for cleft palate, using the Sommerlad technique, at age 6 months, when compared to the surgery using the Sommerlad technique at age 12 months, improves velopharyngeal function at age 5 years.

The null hypothesis for the primary outcome is that there is no difference in velopharyngeal function at 5 years of age, between those that had surgery at 6 months of age compared to surgery at 12 months of age for cleft palate, using the Sommerlad technique. The alternative hypothesis is that there is a difference between the two groups.

The primary outcome for the TOPS trial is insufficient velopharyngeal function at 5 years of age.

Secondary research objectives are to determine whether surgery for cleft palate, using the Sommerlad technique, at age 6 months, when compared to surgery at age 12 months, improves any of the following outcome measures:

1. Velopharyngeal function at age 5 years;
    a. Velopharyngeal composite score summary (VPC-sum)
    b. Insufficient velopharyngeal function (VPC-rate)
2. Velopharyngeal function at age 3 years;
    a. Insufficient velopharyngeal function (VPC-rate)
    b. Velopharyngeal insufficiency symptoms
3. Canonical babbling at age 12 months:
    a. Canonical babbling present
    b. Canonical babbling ratio
    c. Consonant inventory
4. Articulation at age 3 years:
    a. Percent consonants correct (PCC)
    b. Percent correct placement (PCP)
    c. Percent correct manner (PCM)
    d. Non-oral consonant errors
    e. Oral consonant errors
5. Articulation at age 5 years:
    a. Percent consonants correct (PCC)
    b. Percent correct placement (PCP)
    c. Percent correct manner (PCM)
    d. Non-oral consonant errors
    e. Oral consonant errors
6. Postoperative/long term complications:
    a. Dehiscence
    b. Infection
    c. Evidence of fistula
7. Hearing level:
        a. At 12 months
            i. Abnormal Transient Otoacoustic Emission (TEOAE)
            ii. Abnormal Sound field audiometry
        b. At 3 and 5 years
            i. Abnormal Pure tone audiometry in at least one ear
            ii. Abnormal Pure tone audiometry in both ears
            iii. Severity of better ear (normal, mild, moderate, severe, profound)
8. Middle ear function
    a. Flat line Tympanogram in at least one ear (12 months, 3 years, 5 years)
    b. Flat line Tympanogram in both ears (12 months, 3 years, 5 years)

9. Dentofacial development at age 5 years:
    a. Soft tissue ANB (the angle between soft tissue nasion, A point, and B point on a profile photograph [23])
    b. Maxillary arch constriction score (using modified Huddart/Bodenham scoring system [24,25])
10. Growth at 12 months:
    a. Nude weight
    b. Crown to heel length
    c. Occipitofrontal circumference

8 Investigational plan and study design

8.1 Description of overall study design and plan

TOPS is an international, multi-centre trial using a parallel arm design aiming to detect whether surgery at 6 months is superior to surgery at 12 months.

Infants with a diagnosis of cleft palate are randomised to receive primary surgery for cleft palate using a standardised technique (the Sommerlad technique [26]) at either age 6 months or 12 months (corrected for gestational age).

8.2 Treatments studied

All participants recruited to the trial will receive primary surgical repair of cleft palate using the Sommerlad technique.

Each surgeon, including those who already use the Sommerlad technique, will be provided with written descriptions of the surgical procedures and a video prepared by Mr Sommerlad illustrating all steps to be followed in the surgery. Each surgeon will also attend surgical calibration where they will receive instruction by Mr Sommerlad. The written descriptions of the surgical technique have been reproduced from B.C. Sommerlad's publications [27-28]. Each participating surgeon will receive a copy of these four weeks prior to calibration. The procedures are also described in supplementary trial specific document #15, an appendix to the TOPS Protocol.

Participants will be randomised to either surgery at 6 months or surgery at 12 months of age (corrected for gestational age).

8.3 Treatment compliance

The randomisation details for a patient are provided in *Form 1: Screening and Randomisation* which includes:

- Randomised treatment;
- Estimated date of surgery;
- Randomisation date and number.

The actual date of surgery is also collected on *Form 2: Surgery and Postoperative Care.* Acceptable timescales for surgery are provided in *Section 8.2* of the TOPS Protocol, the acceptable window (earliest and latest date) for surgery will be calculated using the estimated surgery date on *Form 1: Randomisation and screening*.

In this trial of surgical timing, a summary of compliance with study timing will take consideration of:

- Number of withdrawals prior to surgery and the reasons for withdrawal [*Form W1: Withdrawal*];
- Days deviated from scheduled surgery window and the reasons for deviatiation.

8.4 Patient population studied

The trial population is up to 650 children from 23 centres within the UK, Brazil and Scandinavia. Infants with cleft palate who are considered medically fit for operation at 6 months, corrected for gestational age, will be included in the trial.

Gestational age is the length of pregnacy, which is typically 40 weeks. For surgery and 12 month visits a gestational age correction based on the 40 weeks term will be used. Therefore the date of birth will be corrected to 40 weeks gestation with 6 and 12 month visits estimated from this correction. Where gestational age is missing (from *Form 1 and Form C1)* it will be assumed to be 40 weeks (full-term) when performing the date of birth correction for the 6 month and 12 month visits.

8.4.1 Inclusion criteria

The following are inclusion criteria in the TOPS trial:

a. Infants with isolated cleft palate;
b. Medically fit for operation at 6 months, corrected for gestational age;
c. Written informed proxy consent;

d. One parent/carer a native language speaker of the majority language in the country of residence.

8.4.2 Exclusion criteria

The following are exclusion criteria in the TOPS trial:

a. Consent not obtained;
b. Infants with syndromic cleft palate (except Van der Woude syndrome, which can be included if hearing is not affected) or severe developmental delay;
c. Congenital sensorineural hearing loss or structural middle ear anomalies;
d. Variation in the anatomical presentation is such that the surgeon who will perform the procedure considers that one stage closure with the Sommerlad technique would be inappropriate;
e. Submucous cleft palate (defined by the classical triad of signs: bifid uvula, bony defect of the hard palate and muscular diastasis)[29];
f. Where the language spoken at home is not the majority language in the country of residence.

8.4.3 Removal of patients from therapy or assessment

If a decision is made to withdraw from the allocated timing of surgery then the infant may have a different timing of surgery to that allocated but will remain in the trial and have scheduled follow ups should the parent/legal guardian be happy to do so.

Form W1: Withdrawal will collect details on infants who have withdrawn and the reasons for withdrawal. Data from all withdrawn infants will be assessed as described in *Table 1*.

Table 1: Approaches to handling withdrawals from the trial

Level of withdrawal	Decision to withdraw	Action	Data analysed
Complete withdrawal – no follow up, no consent to use collected data	Parent/ Guardian/ Clinician	Removed from all assessments.	Only randomisation number, date and information collected on *Form W1: Withdrawal* will be analysed.
No follow up, consent to use collected data	Parent/ Guardian	Removed from all assessments from the date of withdrawal	Data collected up to the point of withdrawal date, and no data collected from the point of withdrawal, will be analysed.

Level of withdrawal	Decision to withdraw	Action	Data analysed
	Clinician	Analyse only data collected.	Data collected as part of the main trial will be analysed In some cases participants were invited back for a 5 year visit. Those that attended this visit will have the 5 year visit data analysed.
Continue follow up	Parent/ Guardian/ Clinician	Patient not removed from assessments.	Data collected as part of the main trial will be analysed.
Complete withdrawal, but later reconsented	Parent/ Guardian/ Clinican	Removed from all assessments within withdrawal time period.	Remove data for withdrawal time period, include data collected up to first withdrawal date and after reconsent date.

8.5 Consent process

The consent process is described in *Section 12* of the TOPS Protocol.

8.6 Blinding

The nature of the interventions prevents this trial from being blind to participants or their carers. However, speech outcomes are assessed as far as practically possible by blinded assessors. Case report forms (CRFs) do not disclose randomisation arm as far as possible.

Allocation of treatments was concealed using a minimisation method via a web randomisation system for a randomisation schedule which made the next allocation difficult to determine.

8.7 Method of assignment to treatment

TOPS used a secure online randomisation system with a minimisation program which assigned the next patient to the smaller treatment group with a 0.8 probability, or if groups are balanced, with 0.5 probability. The minimisation program was applied within strata defined by surgeon and size of cleft (soft palate only or soft and hard palate).

8.8 Sequence and duration of all study periods

A schematic of the study design can be found in *Section 1,* and a table of trial assessments can be found in *Section 9,* of the TOPS Protocol.

8.9 Schedule of assessments

The full schedule of trial assessments and timeline of data collection can be found in *Section 9* of the TOPS Protocol.

9 Listing of outcomes

The TOPS primary and secondary outcomes are listed below. Definitions and derivations of each outcome are provided in *Section 18.6.*

9.1 Primary outcome

The primary outcome is insufficient velopharyngeal function at age 5 years.

This is a dichotomous outcome of whether the child has insufficient function or not. This is measured by velopharyngeal composite score summary (VPC-sum) which is an overall score on the scale 0-6. Insufficient velopharyngeal function is determined by a score of 4 or more on the VPC-sum scale.

9.2 Secondary outcomes

The secondary endpoints are as follows:

1. Velopharyngeal function at age 5 years [a]:
    a. Velopharyngeal composite score summary (VPC-sum) [b, c]
    b. Insufficient velopharyngeal function (VPC-rate) [d]
2. Velopharyngeal function at age 3 years [a]:
    a. Insufficient velopharyngeal function (VPC-rate) [e]
    b. Velopharyngeal insufficiency symptoms [b]
3. Canonical babbling at age 12 months [a, f]:
    a. Canonical babbling present [g]
    b. Canonical babbling ratio [g]
    c. Consonant inventory [g]
4. Articulation at age 3 years [a]:
    a. Percent consonants correct (PCC) [b]
    b. Percent correct placement (PCP) [b]
    c. Percent correct manner (PCM) [b]

   d. Non-oral consonant errors [b]
   e. Oral consonant errors [b]
5. Articulation at age 5 years [a]:
   a. Percent consonants correct (PCC) [b]
   b. Percent correct placement (PCP) [b]
   c. Percent correct manner (PCM) [b]
   d. Non-oral consonant errors [b]
   e. Oral consonant errors [b]
6. Postoperative/long term complications [h]:
   a. Dehiscence
   b. Infection
   c. Evidence of fistula
7. Hearing level:
   a. At 12 months [f]:
      i. Abnormal Transient Otoacoustic Emission (TEOAE)
      ii. Abnormal Sound field audiometry
   b. At 3 and 5 years [i]:
      i. Abnormal Pure tone audiometry in at least one ear
      ii. Abnormal Pure tone audiometry in both ears
      iii. Severity of better ear (normal, mild, moderate, severe, profound)
8. Middle ear function:
   a. Flat line Tympanogram in at least one ear (12 months [f], 3 years, 5 years)
   b. Flat line Tympanogram in both ears (12 months [f], 3 years, 5 years)
9. Dentofacial development at age 5 years:
   a. Soft tissue ANB (the angle between soft tissue nasion, A point, and B point on a profile photograph)
   b. Maxillary arch constriction score (using modified Huddart/Bodenham scoring system)
10. Growth at 12 months [f]:
   a. Nude weight
   b. Crown to heel length
   c. Occipitofrontal circumference

Notes

a The methods involved in obtaining speech recordings vary according to the child's age. Importantly velopharyngeal function at 3 years uses an age appropriate alternative to the VPC-sum used at 5 years

b Source of speech endpoint: Pictures (Phonetics/Transcriptions)

c Source of speech endpoint: 9 word string

d Source of speech endpoint: Five year spontaneous speech / retelling of bus story

e Source of speech endpoint: Three year spontaneous speech

f Timing of all 12 months assessments are adjusted for gestational age

g Source of speech endpoint: Video of play interaction

h The time frame for post-operative/long term complications is as follows:

- *Dehiscence:* Dehiscence is assessed 48 hours and 30 days postoperatively. Secondary surgery with dehiscence as the reason for this is also assessed.
- *Infection:* Infection is assessed 48 hours and 30 days postoperatively.
- *Fistula*: Fistula is assessed 30 days post operatively and at the 3 and 5 year follow up.

i If Pure tone audiometry could not be performed then Sound field audiometry will be analysed, if measured.

10 Determination of sample size

The sample size calculation was based on a test for proportions using a normal approximation: 292 participants per arm will allow a reduction in insufficient velopharyngeal function at five years from 40% to 29% to be detected with 80% power using a chi-square test (2 sided significance test at 0.05 level). The estimate of 40% was obtained from the pilot trial in 50 five year old participants, collected during the planning period for the grant application. To allow an approximate drop out of 10%, 648 participants will be recruited.

In this trial, restating the power for 300 participants per arm allows for a reduction in insufficient velopharyngeal function at five years from 40% to 29% to be detected with 81% power using a chi-square test (2 sided significance test at 0.05 level).

To consider the potential impact of variability around the value of 40%, 300 participants per group provides 80% power to detect a reduction from 30% to 20% and 76% power to detect a reduction from 20% to 12%.

11 Study framework

The overall objectives for each of the study outcomes (primary and secondary) is to test the superiority of surgery, using the Sommerlad technique, at 6 months as opposed to surgery at 12 months for cleft palate infants.

12 Confidence intervals, p-values and multiplicity

A p-value of 0.05 or less will be used to declare statistical significance for all analyses, p-values will be reported to two significant figures. Rather than adjust for multiplicity, relevant results from other studies already reported in the literature will be taken into account in the interpretation of results. Percentages will be presented to one decimal place and continuous summary statistics will be given to maximum of two decimal places.

13 Timing and objectives of interim and final analyses

13.1 Interim monitoring and analyses

Details on interim analyses are compatible with those found in *Section 10.5* of the TOPS Protocol.

The trial is monitored by an independent Data and Safety Monitoring Board (DSMB)[30] which assesses the trial data and takes into account the current world-wide evidence. DSMB members comply with a trial-specific DSMB charter according to the International Conference on Harmonisation Good Clinical Practice (ICH GCP)[31] and the National Institute for Dental and Craniofacial Research guidelines[32].

The initial analysis of trial data for DSMB review was planned, following the completion of 20 surgical procedures, to assess recruitment rates and adverse events (AEs). Timings of subsequent analyses were determined on the basis of recruitment rates and anticipated that these would be approximately every 12 months. The DSMB monitor adherence to the TOPS Protocol and the quality of the accruing data. The DSMB could request additional interim analyses triggered by concern regarding reported serious adverse events (SAEs).

13.2 Interim analysis at the end of the initial funding period

An interim analysis was planned to coincide with the end of the initial funding period. By this time point 300 participants were expected to have reached the 3 year assessment providing 51% power to detect a reduction in insufficient velopharyngeal function at 3 years from 40% to 29% using a chi-square test with statistical significance at the 0.05 level. However, recruitment to the study was slower than expected and this impacted the value of the planned interim analysis in which 300 participants had been expected to contribute.

Due to the slower rate of recruitment, the expected number of infants at 3 year assessment by the end of June 2014 was 33, with 142 participants by June 2015, if a 12 month no cost extension was provided. The assessment of velopharyngeal function at age 3 years was intended to be a proxy for the primary outcome assessment at age 5 years with the purpose of the interim analysis being to inform the next funding period. This interim analysis was therefore not performed due to the low numbers of infants at the timepoint required. Despite the interim analysis not being undertaken, trial monitoring and review by the DSMB continued.

Following each DSMB meeting, the DSMB will make recommendations to the Trial Steering Committee (TSC) as to the continuation of the trial. The trial could be ended, or recommended to continue, as appropriate. Importantly, statistical considerations alone are not adequate to determine the end of the trial due to the over-emphasis placed on the p-value resulting from hypothesis tests. Clinical judgment is essential to the process to account for unexpected AEs and balance issues of safety and efficacy in light of any new external information. The decision to stop recruitment should depend on whether the results will be convincing to the medical community.

13.3 Final analysis

All outcomes will be analysed after the end of the trial, which is defined in *Section 9.2* of the TOPS Protocol as “the date of database lock”. This is the date on which data modification privileges are withdrawn from the trial database.

14 Disposition of participants

14.1 Screening, eligibility and recruitment

A CONSORT[33] diagram will summarise the number of participants who were:

1. Assessed for eligibility
2. Not randomised, where all participants not randomised will be accounted for within the following mutually exclusive sub groups
    i. Not eligible
    ii. Consent not provided
    iii. Consent provided though not randomised
3. Randomised
4. Received the randomised allocation
5. Did not receive the randomised allocation, where all participants not receiving randomised allocation will be accounted for within the following mutually exclusive sub groups

   i. Surgery date outside acceptable time window
   ii. Other (give reasons)
6. Withdrawn from follow up
   i. No follow up, no consent to use collected data
   ii. No follow up, consent to use collected data
7. Intention-to-treat (ITT) population
8. Included in primary outcome set
9. Not included in primary outcome set (give reasons)

Reasons for non randomisation will be summarised in a table with the following categories for reasons, within the subgroup as listed:

1. Not eligible (NE)
   i. Infant is not medically fit for operation at age 6 months corrected for gestational age
   ii. Neither parent or carer a native language speaker of the majority language in the country of residence
   iii. Known associated syndrome or severe developmental delay
   iv. Congenital sensorineural hearing loss or structural middle ear anomalies
   v. Cleft too wide for closure with Sommerlad technique
   vi. Other variation in the anatomical presentation is such that one stage closure with the Sommerlad technique would be inappropriate
   vii. Submucous cleft palate
   viii. Language spoken at home is not the majority language of the country of residence
   ix. Unable to complete clinical genetics assessment prior to 6 month admission age
   x. Unsuitable (reasons specified categorised)
   xi. At least one parent/carer is not a native language speaker of the majority language*
   xii. Clinical genetics assessment not completed prior to 6 month admission age*
2. Consent not provided (CNP)
   i. Preference for 6 months timing of surgery
   ii. Preference for 12 months timing of surgery
   iii. Other reason (reasons specified categorised)
   iv. Parents/Carer declined participation: No reason given

*Not eligible reason on superseded version of TOPS Screening Log.

Frequencies of each reason provided will be presented along with percentages using the number of participants within that category (total number of participants NE and CNP respectively as the denominator).

Note that other reasons specified by free text will be reclassified with clinical input where required for each category above.

The following centre specific randomisation details will be reported. This data will also be summarised for all participants across all centres:

1. Date centre opened
2. Date centre closed
3. Planned total recruitment
4. Date of first randomisation
5. Date of last randomisation
6. Total number randomised.

The following centre specific screening details will be reported. These data will also be summarised for all participants across all centres. Frequencies and percentages will be reported:

1. Number of participants screened (s)
2. Number of participants deemed ineligible (s_i, denominator s)
3. Number of participants who did not provide consent (s_{ii}, denominator s)
4. Number of participants who consented but were not randomised (s_{iii}, denominator s)
5. Number of participants who were randomised (s_{iv}, denominator s)
6. Consent rate (reported as percentage, calculated $\frac{s_{iii}+s_{iv}}{s_{ii}+s_{iii}+s_{iv}}$)

14.2 Post randomisation discontinuations

The timing of withdrawal in relation to surgery and scheduled visits, level of withdrawal, who made the decision and reason for withdrawal will be summarised both overall and for each randomised group. Frequencies will be presented along with percentages using the number of participants who withdrew as the denominator. Note, other reasons specified by free text will be reclassified with clinical input where required.

The number lost to follow up both overall and within each randomised group will be reported and the reasons where known will be documented. Any deaths and their causes will be reported separately.

Reasons for participants not receiving the randomised allocation will be summarised in a table with categories dependent on the free text provided on the CRF. Frequencies will be presented along with percentages using the number of participants who were randomised as the denominator both overall and for each randomised group.

Withdrawals from the trial will be presented as line listings detailing:

- Treatment arm
- Who made the decision to withdraw participant from trial:
    - Clinician
    - Parent/Guardian
    - Clinician and participant
- Level of withdrawal
    - Continue follow up
    - No follow up, consent to use collected data
    - No follow up, no consent to use collected data (complete withdrawal)
- Reason for withdrawal;
    - Identification of syndrome
    - Identification of a severe developmental delay
    - Other (free text)
    - Reason not given
- Patient later reconsented, where yes:
    - Follow up visits missed due to withdrawal
    - Point of reconcent.

15 Protocol deviations

Protocol deviations that will be reported are defined in the monitoring plan for the trial, see Appendix *Table 22*. Due to the nature of the intervention, adherence to treatment i.e. timing of surgery cannot be classified prior to unblinding however all other protocol deviations will be classified prior to unblinding of treatment to the statistical team. Protocol deviations will be defined and signed off using *ST001TEM04 Protocol deviations and data set definitions* template prior to unblinding as far as practically possible. A summary of the analysis populations will be provided which include: Randomised, Safety, Intention-to-treat, primary outcome analysis and per-protocol analysis populations.

The number (and percentage) of participants with each separate protocol deviation will be presented in this analysis report along with the number (and percentage) of participants with (i) at least one major protocol deviation; (ii) at least one minor protocol deviation; and (iii) at

least one protocol deviation of any classification (minor or major). An overall summary and per patient protocol deviations will be provided. No formal statistical testing will be undertaken.

Treatment and pre-specified follow up visits occurring outside the visit windows will also be summarised as protocol deviations. The treatment and follow up visit windows are defined in *Section 8.2* and *Section 9.1.4* of the TOPS protocol.

In addition to the protocol deviations, a check that the randomisation process was balanced will be presented in the statistical analysis report. This check will be performed by summarising the number of participants within each stratification factor, split by treatment group. This check will be summarised as frequencies, with percentages within treatment group presented, where percentages are calculated using the denominator of the number of participants within that strata. In addition, any randomisation errors, such as the duplication of treatment allocations or gaps in sequential numbers within centres will be presented as line listings and any issues described.

16 Unblinding

Not applicable.

17 Completeness of follow up assessments

Adherence to the time points (30 days, 12 months, 3 years and 5 years of age) will be classified as at least one assessment visit completed in the expected window*. Where no assessments are reported within the window, the assessment visit closest to the visit window will be used. Adherence to each visit will be summarised separately as (i) assessed within window; (ii) assessed outside window; and (iii) not assessed. Windows for follow up timepoints defined within the TOPS Protocol are provided in *Table 2*. This will be reported overall, by site and treatment group.

* Assessments can consist of any of the following: case report form (CRFs); speech recording; photograph; and/or impression taken, depending on timpoint, with *Section 17.1*, *Section 17.2*, *Section 17.3*, and *Section 17.4* providing further detail on these respectively.

Table 2: Windows for follow up visits according to the protocol

Time point	Estimated visit date	Early window limit	Late window limit
12 months [1]	*6 months surgery* [2, 3]*:*		
	Estimated visit date is 52 weeks from the infant's	-4 weeks of estimated visit date.	+4 weeks of estimated visit date.

Time point	Estimated visit date	Early window limit	Late window limit
	date of birth corrected for gestational age.		
	12 month surgery [3, 4]*:*		
	52 weeks from the infant's date of birth corrected for gestational age.	Visit window same as surgery: -2 weeks of estimated date	all assessments to take place prior to surgery: +4 weeks of estimated date
3 years [5]	156 weeks from infants date of birth	-4 weeks of estimated visit date	+4 weeks of estimated visit date
5 years [6]	260 weeks from date of birth	-4 weeks of estimated visit date	+4 weeks of estimated visit date

[1] When correcting for gestational age, full term defined as day 1 of the 40th week of pregnancy.

[2] Speech assessments should be completed between 48 weeks and 18 months of age (equivalent to 78 weeks).

[3] Audiology and fistula assessments should not be completed after 23 months of age (equivalent to 100 weeks).

[4] Speech assessments should be completed before surgery (between 50 weeks of age and surgery date).

[5] Participants should not be recalled after 3 years 11 months of age, acceptable speech window is 152 weeks from date of birth to 3 years 11 months of age (equivalent to 204 weeks).

[6] Participants should not be recalled after 5 years 11 months of age, acceptable speech window is 256 weeks from date of birth to 5 years 11 months of age (equivalent to 308 weeks).

17.1 Case report forms

The completeness of case report forms (CRFs) will be reported at 12 months, 3 years and 5 years of age. *Table 3* indicates which CRFs are to be completed at each timepoint.

Table 3: Case report form by follow up visit

CRF	Screening and randomisation	Surgery	12 months[1]	3 years[1]	5 years[1]
Form S1: Screening log	X				
Form C1: Genetics assessment	X				
Form 1: Screening and randomisation	X				
Form 2: Surgery and postoperative care[2]		X			
Form C2: Blood sample for genetic assessment		X			
Form 3: Late Complications[3]		X			
Form 4: 12 month speech assessment			X		
Form 5: 12 month growth assessment			X		
Form 6: 12 months audiology data collection			X		
Form 7: Fistula assessment			X	X	X
Form 8: Audiology data collection 3 and 5 years				X	X
Form 9: 3 Year Speech Assessment				X	
Form 10: Denver II assessment at 3 years				X	
Form 11: 3 Year speech therapy questionnaire				X	
Form 13: 5 Year Speech assessment					X
Form 14: 5 Year speech therapy questionnaire					X
Inteligibility in Context Scale					X
Form T1: Dental/ Maxillary arch sample transfer		X			X
Form T2: Log for transfer of speech recordings or participant photographs		X	X	X	X
Form T3: Speech recording sample transfer			X	X	X
Form A1: Serious Adverse Events	Complete for each SAE in first 30 days following surgery				
Form M1: Concurrent medication	Complete for each concurrent medication during and for 30 days post surgery				
Form U1: Unanticipated Problems	Complete for each unanticipated problem for duration of the trial				
Form 12: Secondary Surgery	Complete for each participant receiving additional surgery during trial				
Form 15: Participant Transfer	Complete for each participant transferred from one TOPS site to another TOPS site during the trial (doesn't apply to Bauru site)				
Form W1: Withdrawal	Complete for each participant withdrawal during the trial				

[1]Time windows for follow up visits are detailed in *Table 2*.

[2]Post-operative care is described up to 48 hours post surgery. Surgery should occur 2 weeks before to 4 weeks after the target date.

[3]Late Complications should be recorded up to 30 days post surgery.

17.2 Speech recordings

The types of speech recordings taken by timepoint, and used for outcome assessment, are summarised in *Table 4*.

Appendix Table 23 summarises the speech outcomes collected by timepoint, mapped against the speech recording type.

Table 4: Summary of recording types by timepoint

Timepoint	Speech sample	Recording ID	Recording type
12 months	Spontaneous speech (Play interaction)	12m_1	Video
3 years	Phonetics[1]	3y_1	Video
	Spontaneous speech[2]	3y_2	Video
5 years	Phonetics[1]	5y_1	Video
	Spontaneous speech[2]	5y_2	Video[3]
	Nine-word string	5y_3	Audio
	Bus story	5y_4	Video[3]

Notes:

(1) Phonetics also known as naming test.

(2) Spontaneous speech also known as rating.

(3) Spontaneous speech and bus story are taken as a video during the speech visit, due to remote assessment the video files were converted to audio files for the assessments.

Where the primary recording type is video (*Table 4*), audio files may be available instead of video and will be included for assessment wherever possible.

Form 4, Form 9 and *Form 13* will indicate if recordings were taken at timepoints 12 months, 3 years and 5 years of age, respectively. The data on these forms will be cross checked against a master file held by the Administrative Centre (AC), who, in collaboration with the Core Speech Group (CSG), conduct quality checks on recordings received. Any discrepancies between lists will be queried and, once settled, assessment tables presented using confirmed total received.

Table 5 presents reasons for not completing a speech recording for each timepoint, as recorded on *Form 4, Form 9* and *Form 13* respectively. The same question for non completion was proposed if:

- Less than 10 target sounds were named and/or the child produced less than 15 minutes of spontaneous speech at 3 years of age
- Less than 25 target sounds were named and/or the child produced less than 5 minutes of spontaneous speech at 5 years of age.

Table 5: Reasons for not completing a speech recording by timepoint

Timepoint	Reason
12 months	• Child sleeping • Child shy/quiet • Child crying • Equipment failure • Other
3 years and 5 years	• Non-cooperative • Child has a developmental delay • Child unable to name pictures in native language • Equipment failure • Other

When not taken, a second attempt at a recording was made wherever possible, the CRF at each timepoint indicates reattempts.

At 12 months, 3 years and 5 years the reasons for non-completion of the primary recording will be reported with, where appropriate, the number of rescheduled recordings, and whether or not a recording was subsequently completed or not (see *Table 6*), as appropriate.

Table 6: Reasons for no further attempt to complete the speech recording

Timepoint	Reason
12 months	• Unable to complete before surgery • Parents unable to make return journey • Other
3 years and 5 years	• Parents unable to make a return journey • Child has a developmental delay and inappropriate to invite back for a repeat visit at 3/5 years • Child has a syndrome and inappropriate to invite back for a repeat visit at 3/5 years • Other

Speech recordings are assessed for quality before being analysed for outcomes by TOPS team members at the AC and CSG. All recordings are classified as include or exclude based on a strict, outcome dependent, rule set, determined by the CSG. Recordings classified as exclude will not be used in the main analysis of speech outcomes and reasons for exclusions will be reported.

17.3 Photographs

Photographs are taken at surgery and at 5 years of age for each child in the study.

At surgery, a photograph of the child's unrepaired palate is taken however, this is not an outcome. Photographs are received at the Data Co-ordinating Centre (DCC) who hold a master list of surgical photographs received.

At five years, a photograph of the child's face profile (frontal and right/left lateral) is taken. The left lateral photograph is used to measure dentofacial development (*Outcome ID: SO_9a_5y*; *Appendix Table 23*), where the left is not provided or fails quality checks but the right lateral photograph is provided and passes quality checks the right photograph will be used . The AC conduct quality checks on collected photographs to ensure the outcome can be derived accurately. The master file held by the AC will be used to determine the number of photographs received and of adequate quality.

17.4 Impressions

Impressions are taken at surgery and at 5 years of age for each child in the study.

At surgery, a maxillary arch impression is taken and is not an outcome. Maxillary arch impressions at surgery and dental impressions at 5 years are received at the AC who hold a master list of impressions received.

At five years, dental impressions are taken and are used to calculate the Maxillary arch constriction score (*Outcome ID: SO_9b_5y, Appendix Table 23*). The AC conduct quality checks on collected dental impressions to ensure the outcome can be derived accurately. The master file held by the AC will be used to determine the number of dental impressions received and of adequate quality.

18 Efficacy evaluations

18.1 Data sets analysed

The principle of intention-to-treat (ITT), as far as practically possible, will be the main strategy of the analysis adopted for the primary outcome and all the secondary outcomes. These analyses will be conducted on all randomised participants, in the group to which they were allocated, and for whom the outcome(s) of interest have been observed/measured. Unless specified otherwise in the analysis section, no imputations will be made.

The membership of the analysis set for each outcome will be determined and documented and reasons for participant exclusion will be given. Reasons may include missing data or patient withdrawal.

The per protocol population will mirror the ITT population but exclude any participants defined as having a major protocol deviation (see *Section 15*) as specified within the monitoring plan. The per protocol analysis will only be considered in the event of major protocol deviations in more than 10% of the ITT analysis population and apply to a secondary analysis of the primary outcome only.

The numbers within the ITT data set will be presented as count data and a percentage using the total randomised as a denominator. This will be presented both overall and split by treatment group.

The safety dataset will classify participants who have surgery before 9 months of gestational corrected age as received 6 months surgery, and surgery at 9 months of gestational corrected age or beyond as received surgery at 12 months.

Update from blind review:

A per protocol analysis will not be performed, instead two sensitivity analysis will be performed whereby two major protocol deviations are removed, see Section 18.6.1.3.3.

18.2 Demographic and other baseline characteristics

The following baseline characteristics will be reported.

Patient demographic characteristics [*Form 1: Screening and Randomisation*] ;

- Gender
    - Male
    - Female
- Gestational age (weeks)
    - Gestational age is collected on *Form 1* and *Form C1*. This data will be summarised from *Form C1*, and where missing this will be taken from *Form 1.*
- Size of cleft
    - Soft palate only
    - Soft and hard palate

Clinical characteristics at baseline [*Form C1: Genetic Assessment*];

- Ethnicity
    - White
    - Black
    - Asian
    - Chinese
    - Mixed
    - Other
    - Not Stated
- Weight at examination (kg)
- Length at examination (cm)
- Occipitofrontal circumference (cm)
- Interpretation of Denver
    - Normal
    - Suspect
    - Un-testable
- Diagnosis
    - Known syndrome
    - Unknown syndrome

- Severe developmental delay
- Uncertain
- Non-syndromic

Note: *Form C2: Blood Sample for Genetic Assessment* will be used instead of *Form C1* if no results where available at screening.

These will be presented both overall and for each randomised group. Clinical genetics will be described by region (UK, Scandinavia, Brazil).

Racial categories will be reported overall and split by gender, in accordance with the NIH categories. The ethnicity categories [*Form C1: Genetic Assessment*] will be used to categorise children into each of the following racial categories:

- White (as reported on *Form C1*)
- Black (Black or African American ethnicity)
- Asian (as reported on *Form C1*)
- Chinese (as reported on *Form C1*)
- Mixed (More than one ethnicity provided on *Form C1*)
- Other (classified using ethnicity freetext for other)
- Not Stated

Categorical data will be summarised by frequencies and percentages. Continuous data will be presented by means and standard deviations (SDs), or medians and inter-quartile range (IQR) if data are skewed. Minimum and maximum values will also be presented for continuous data. Tests of statistical significance will not be undertaken for baseline characteristics; rather the clinical importance of any imbalance will be noted. The amount missing in each case will be summarised.

18.3 Compliance with treatment

For details on treatment compliance, see *Section 8.3*.

Treatment adherence will be summarised descriptively as frequency tables, both overall and split by randomised group and centre, presenting the number of participants:

- Receiving the surgery, as allocated, within the scheduled surgery window (two percentages presented: denominator: number received surgery and denominator: number randomised to surgery)

- Receiving the surgery, as allocated, outside the scheduled surgery window. Days deviated from scheduled surgery window, and reasons for deviation, will be provided. (two percentages presented: denominator: number received surgery and denominator: number randomised to surgery)
- Not receiving the surgery as allocated i.e. withdrawal from surgical timing. Reasons will be provided. (denominator: number randomised to surgery)

The number of days early, number of days late and the absolute number of days deviated will be presented. The number of days deviated will be calculated as shown in *Table 7*.

Table 7: Derivation of days deviated from acceptable treatment window

Deviation	Calculation	Notes
Number of days early	Earliest expected date – visit date	Earliest expected date = expected date – 4 weeks
Number of days late	Visit date – latest expected date	Latest expected date= expected date + 4 weeks
Absolute number of days deviated	Number of days outside window	Zero if within window.

The number of participants overall and split by treatment group who deviated (early, late and either) will be presented. The number of days early, late and absolute deviated will be summarised as means and standard deviations (SDs), or medians and inter-quartile range (IQR) if data are skewed. Minimum and maximum values will also be presented.

18.4 Compliance with study follow up and assessments

For details on follow up visit compliance, see *Section 17*. The number of assessments inside the acceptable time window (see *Table 2*) and those not will be reported overall and by treatment group as frequencies and percentages, with the denominators of all those collected for each assessment and randomised group.

For each assessment timepoint below and each outcome in the *outcome* column of *Table 8, Table* 9*, Table 10* and *Table 11* the items under *data collected* will be reported as frequencies and percentages to indicate outcome completion with denominators of randomised group and overall cohort. Where applicable reasons for non-completion will be reported.

Quality of speech recordings, photographs and dental impressions will be reported, see *Section 17*, as frequencies and percentages with denominators of randomised group and overall cohort.

Audiometry at 12 months requires Sound field audiometry (*Outcome ID: SO_7aii_12m; Appendix Table 23*), while 3 and 5 years requires Pure Tone audiometry (*Outcome ID: SO_7bi_3y, SO_7bii_3y , SO_7bi_5y and SO_7bii_5y; Appendix Table 23*). Both of these should be performed at four test frequencies (500Hz, 1000Hz, 2000Hz and 4000Hz), therefore the number of participants with 0 to 4 frequencies tested will be reported.

18.4.1 Post-operative assessments

Table 8 provides the data collection required for the post-operative outcomes at 48 hours and 30 days post surgery.

Table 8: Summary of postoperative complications assessments

Assessment	Outcome – Outcome ID [1]	Data collected	Source
Dehiscence	Yes • SO_6a_48h • SO_6a_30d	• Yes/No • Free text	Form 2 Form 3
Infection	Yes • SO_6b_48h • SO_6b_30d	• Yes/No • Free text	Form 2 Form 3
Fistula	Yes • SO_6c_30d	Fistula Assessment • Evidence • Length (mm) • Maximum width (mm) • Position	Form 3

[1] *Appendix Table 23* provides further detail of outcomes by mapping on Outcome ID

18.4.2 12 month assessments

Table 9 provides the data collection required for the 12 month assessments.

Table 9: Summary of 12 month assessments

Assessment	Outcome – Outcome ID [1]	Data collected	Source
Growth • Nude weight (grams) • Crown to heel Length (cm) • Occipitofrontal Circumference (cm)	Yes • SO_10a_12m • SO_10b_12m • SO_10c_12m	Growth: • Nude weight (grams) • Crown to heel Length (cm) • Occipitofrontal Circumference (cm)	Form 5
Fistula[2]	No	Fistula Assessment • Evidence • Length (mm) • Maximum width (mm) • Position	Form 7
Audiology • TEOAE • Sound field	Yes • SO_7ai_12m • SO_7aii_12m	• TEOAE completed Sound field: • Soud field test • VRA test • Soudfield and VRA test	Form 6
Middle ear function: • Tympanogram	Yes • SO_8a_12m • SO_8b_12m	Tympanogram test in: • At least one ear (left or right) • Both ears (left and right) Ear canal volume • At least one ear (left or right) • Both ears (left and right)	Form 6

Assessment	Outcome – Outcome ID [1]	Data collected	Source
Speech: • Canonical Babbling	Yes • SO_3a_12m • SO_3b_12m • SO_3c_12m	• Spontaneous speech recording	Form 4 Recording ID: 12m_1

[1] *Appendix Table 23* provides further detail of outcomes by mapping on Outcome ID
[2] For infants who have their primary repair at age 6 months only

18.4.3 3 year assessments

Table 10 provides the data collection required for the 3 year assessments.

Table 10: Summary of 3 year assessments

Assessment	Outcome – Outcome ID [1]	Data collected	Source
Denver II	No	Denver II at 3 years	Form 10
Fistula	Yes • SO_6c_3y	Fistula Assessment • Evidence • Length (mm) • Maximum width (mm) • Position	Form 7
Audiology • Pure tone audiometry • Severity of better ear	Yes • SO_7bi_3y • SO_7bii_3y • SO_7biii_3y	Pure tone audiometry (air conduction) in: • At least one ear (left or right) • Both ears (left and right)	Form 8
Middle ear function: • Tympanogram	Yes • SO_8a_3y • SO_8b_3y	Tympanogram test in: • At least one ear (left or right) • Both ears (left and right) Ear canal Volume • At least one ear (left or right) • Both ears (left and right)	Form 8
Speech: • VPC-rate • Articulation	Yes • SO_2a_3y • SO_2b_3y • SO_4a_3y • SO_4b_3y • SO_4c_3y • SO_4d_3y • SO_4e_3y	• Spontaneous Speech recording (VPC) • Phonetics speech recording	Form 9 Recording ID: 3y_1 and 3y_2

[1] *Appendix Table 23* provides further detail of outcomes by mapping on Outcome ID

18.4.4 5 year assessments

Table 11 provides the data collection required for the 5 year assessments.

Table 11: Summary of 5 year assessments

Assessment	Outcome – Outcome ID [1]	Data collected	Source
Dentofacial development: • Soft tissue ANB • Maxillary arch constriction score	Yes • SO_9a_5y • SO_9b_5y	• Profile Photographs • Dental impressions	Form T1 Form T2 Form 16 Form 17
Fistula	Yes • SO_6c_5y	Fistula Assessment • Evidence • Length (mm) • Maximum width (mm) • Position	Form 7
Audiology • Pure tone audiometry • Severity in better ear	Yes • SO_7bi_5y • SO_7bii_5y • SO_7biii_5y	Pure tone audiometry (air conduction) in: • At least one ear (left or right) • Both ears (left and right)	Form 8
Middle ear function: • Tympanogram	Yes • SO_8a_5y • SO_8b_5y	Tympanogram test in: • At least one ear (left or right) • Both ears (left and right) Ear canal Volume • At least one ear (left or right) • Both ears (left and right)	Form 8
Speech: • VPC-sum[2] • VPC-rate	Yes • PO_5y • SO_1a_5y • SO_1b_5y	• Phonetics speech recording • Spontaneous Speech/ bus	Form 13 Recording ID: 5y_1, 5y_2 and 5y_3

Assessment	Outcome – Outcome ID [1]	Data collected	Source
• Articulation	• SO_5a_5y • SO_5b_5y • SO_5c_5y • SO_5d_5y • SO_5e_5y	story recording (VPC-rate)	

[1] *Appendix Table 23* provides further detail of outcomes by mapping on Outcome ID
[2] Only participants with both speech assessments completed (Phonetics and Hypernasality) will be used in the analysis of the VPC-sum.

18.5 Software for speech outcome assessment

As part of the trial, two bespoke softwares have been developed to facilitate speech outcome assessment by SLTs in the TOPS trial:

- TimeStamper for 12 month speech recordings
- SPEAK for recordings taken at 3 and 5 years.

18.5.1 TimeStamper

TimeStamper allows assessors (SLTs) to review the spontaneous speech recording taken for each participant at 12 months and classify: canonical babbling present, canonical babbling ratio and the consonant inventory (*Outcome ID: SO_3a_12m, SO_3b_12m and SO_3c_12m; Appendix Table 23)*.

The spontaneous speech recording is split into two parts: a and b. Classifications of each outcome will be made by an assessor at the end of each part seperatley. The classifications from parts a and b by an assessor will be combined. Further outcome specific details are provided in *Section 18.6.2.3.1*.

Three assessors will assess each recording and the majority classification used. Every effort will be made to ensure that three complete assessments are made for each participant. Where an assessment is incomplete e.g. part a or part b only assessed, then the classification on the part completed will be used. Where an assessor has made the same assessment twice, i.e. an intra assessment has been made, the classification on the most recent assessment will be used. Due to limitations in the assessment process, some children may have more than three assessors, in this case the first three complete assessments will be analysed.

18.5.1.1 Canonical babbling present

Each spontaneous speech recording comprises a number of syllables produced by the participant, each syllable will be deemed as either “canonical” or “not canonical” by the

assessor using the software keypad. The software exports the total number of "canonical" and "not canonical" syllables, *Kanonisk Stavelse Count* and *Stavelse Count* respectively, for both parts of the assessment (see example output in *Appendix Table 28)*.

The exported data for both parts will be summed to calculate the total syllables "canonical", "not canonical" and the canonical or not canonical as the total number of syllables produced by the participant according to that assessment.

Canonical babbling ratio (CBR) for each assessment will be calculated as the total number of "canonical" syllables divided by the total syllables produced across both parts of the recording, as in the following equation:

$$CBR = \frac{Total\ syllables\ canonical}{Total\ syllables\ canonical\ or\ not\ canonical}.$$

A ratio of at least 15% "canonical" syllables has to be produced in order for a participant to be regarded as being canonical, as agreed with the CSG on 12/09/2019. *Table 12* shows how the majority classification from the three assessor classifications will be determined, as agreed with the CSG on 06/03/2019.

Table 12: Majority classification of canonical babbling through CBR

Majority classification	Number of SLTs per classification	
	Canonical babbling present (CBR ≥ 15%)	Canonical babbling absent (CBR < 15%)
Canonical	3	0
	2	1
Not Canonical	1	2
	0	3

18.5.1.1.1 Alternative assessment of canonical babbling present

An assessor will also classify whether the participant was "Canonical" or "Not Canonical" noted in the column *Sounds* of the software output, example output in *Appendix Table 28*. An alternative definition of the outcome *canonical babbling present* will be determined by each assessor if they classify either part of a participant's recording as "Canonical".

Table 13 shows how the majority classification from the three assessor classifications will be determined.

The alternative classification of *canonical babbling present* was agreed with the CSG on 15/11/2018 and will be used to compare to the formal outcome of *canonical babbling present* for exploratory purposes.

Table 13: Majority classification of canonical babbling

Majority classification	Number of SLTs per classification	
	Canonical	Not Canonical
Canonical	3	0
	2	1
Not Canonical	1	2
	0	3

18.5.1.2 Canonical babbling ratio

The outcome *Canonical babbling ratio (CBR)* is calculated as the total number of "canonical" syllables divided by the total syllables produced across both parts of the recording following the equation in *Section 18.5.1.1*.

Canonical babbling ratio will be defined as the mean ratio of the three independent assessments.

18.5.1.3 Consonant Inventory

As part of an assessment, unique syllables that the participant produces will be noted by the assessor (see example output in *Appendix Table 28)*. Examples of syllables include *ha, ba,* and *ma*. The consonant inventory will comprise of all unique syllables, that are speech like sounds, listed across both parts of the assessment.

The inclusion of speech like sounds only was agreed with the CSG on 04/01/2019. These are sounds which always co-occur with a vowel i.e. as a syllable. The consonant uttered and not the particular vowel is the syllable of interest, meaning that syllables with the same consonant but listed as a unique syllable can be mapped to the consonant of interest only and treated as the same, this mapping structure is provided in *Appendix Table 34*. Sounds noted in the columns titled *Similar* and *Equivalent* will be used in the analysis in the form of the column *Preferred notation*. For example, the syllables *ba, bwa* and *ab* are all provided as equivalent while ʙa and Ba are similar in *Appendix Table 34,* these will all be treated for the analysis under the syllable *ba.*

The listed consonant inventory for a participant is the list of unique consonants identified by at least two of the three assessors. The outcome *consonant inventory* is the number of consonants in the listed consonant inventory.

18.5.2 SPEAK

SPEAK allows an assessor to review the speech recordings taken for each participant at both 3 years and 5 years of age. A version of the software will be developed for age specific assessment, the first at 3 years and the second at 5 years. Only assessments with AssessmentStatus of Complete should be used at 5 years.

18.5.2.1 Phonetic assessments

A phonetic assessment requires a participant to produce a set of specific words predefined by the CSG, in order to assess the production of a particular target sound. The recording shows the participant attempting to produce each word. The target word and target sounds for each language are provided on STSD 17 and 18 for 3 and 5 year speech respectively.

An assessor reviews a participant's attempt at producing each target sound and classifies if that sound is produced. When a participant has failed to produce the target sound, the sound actually produced by the participant in their attempt will be noted. Example output from the software is provided in *Appendix Table 29* and *Table 31* for the 3 years and 5 years assessment respectively.

18.5.2.1.1 3 year mapping from SPEAK output

In the dataset *(Appendix Table 29)* the responses to whether the target sound is correct, "Yes" or "No", are assigned the values 1 and 0 respectively in the column *Correct*. If "No" (0) is selected then the sound that the child uttered is provided in the *Sound Idx* column, while each word can be identified using the *WordIdx* for each language and STSD 17 to identify the target sound. If the column *Notes* states "File Missing" then this word was not available for assessment. If the column *Notes* states "Could Not Assess" or "File Missing" these words should be removed from the child's assessment as outcomes cannot be derived from these assessments.

18.5.2.1.2 3 year phonetics outcome derivation

Table 14 details the derivation of the speech outcomes using the example SPEAK output in *Appendix Table 29* and the 3 year mapping structure provided in *Appendix Table 35.* If the target sound is identified as Incorrect (*Correct = "0", Appendix Table 29*) and the sound identified is the target sound then the Assessor classification of sound produced should be Correct.

Where more than one sound has been provided as an assessment the following rules apply:

- Any sound in *Appendix Table 35* in combination with a glottal plosive (Idx 13) will be treated as a glottal only (Idx 13).
- Any sound in *Appendix Table 35* in combination with Idx 89/ 90/ 91/ 92 (and 93 at 5 years) will be treated as a single sound for all the outcomes defined in *Table 14,* following the *Appendix Table 35* for the sound (not idx 89/ 90/ 91/ 92/ 93). All rules for the VPI symptoms should be followed for the combination of sounds and if a VPI symptom occurs for at least one sound in the combination, then a VPI symptom has occurred.
- Any other sound combinations will be treated as incorrect for consonant, place and manner. Non-oral errors and oral errors can not be derived and therefore, are classified as no error. All rules for the VPI symptoms should be followed for the combination of sounds and if a VPI symptom occurs for at least one sound in the combination, then a VPI symptom has occurred.

If the SLT assessment of the sound is “Incorrect” and the following sounds from *Appendix Table 35* have been provided alone idx 89/ 90/ 91/ 92 (and 93 at 5 years), i.e. not in combination with another sound, then the word assessment should be set to ‘Could Not Assess’.

If the SLT assessment of the sound is “Incorrect” and the following sounds from *Appendix Table 35* and *Appendix Table 36* have been provided, 68/69/ 70/ 71/ 94/ 95/ 96, then these sounds should be recoded as follows:

- Idx 68 (tˢ) to Idx 3 (t)
- Idx 69 (b̥̋) to idx 2|90 (b|˅)
- Idx 70 (d̥̋) to idx 4|90 (d|˅)
- Idx 71 (g̥̋) to idx 10|90 (g|˅)
- Idx 94 (b̥) to idx 2 (b)
- Idx 95 (d̥) to Idx 4 (d)
- Idx 96 (g̊) to Idx 10 (g)

The target sounds with Idx 68/94/95/96 should also be recoded as above.

- For the outcome PCM the manner:
    - ‘Lateral approximant’ should be treated as ‘Approximant’
    - ‘Active Nasal Fricative’ and ‘Lateral, fricative’ should be treated as ‘Fricative’

Table 14: Speech outcome derivation from SPEAK and the mapping structure

Outcome	Assessor classification of sound produced	Comparison of sound produced to target sound	Outcome classification
Correct consonant	Correct	N/A	Correct
	Not correct	Same place, manner and voice as target sound*	Correct
		Different place, manner, or voice to target sound	Not correct
Correct place	Correct	N/A	Correct
	Not correct	Same place as target sound	Correct
		Different place to target sound	Not correct
Correct manner	Correct	N/A	Correct
	Not correct	Same manner as target sound	Correct
		Different manner to target sound	Not correct
Oral consonant error	Correct	N/A	No error
	Not Correct	*Oral consonant error if* column states "n.a" for that Sound	No error
		Target sound is listed in the column *Oral consonant error if* for that Sound	Oral Consonant error
		Oral consonant error if column states "always an Oral Consonant error" for that Sound	Oral Consonant error
Non-oral Consonant error	Correct	N/A	No error
	Not Correct	*Non-Oral Consonant error if* columns states "n.a" for the sound	No error
		Target sound is listed in the column Non-o*ral consonant error if* for that Sound	Non-oral Consonant error
		Non-Oral Consonant error if columns states "always non-oral consonant error" for the sound	Non-oral consonant error

*Idx 51 (ɬ) and 82 (n̥) for PCC should have the manner treated as 'Lateral, fricative' and 'active nasal fricative' and not 'fricative' for the outcome.

Velopharyngeal insufficiency (VPI) symptoms derivation

No VPI symptom occurs if:

i. The target sound is assessed as Correct (*Correct = "1", Appendix Table 29*).
ii. The target sound is identified as Incorrect (*Correct = "0", Appendix Table 29*) and the sound identified is the target sound.

However, if the target sound is assessed as Incorrect (*Correct = "0", Appendix Table 29*) and the sound identified is not the target sound then the following criteria hold, using *Appendix Table 35*:

i. For the sound identified, see *Sound* column, if the *VPI Symptoms present if* column states "always a VPI symptom" then a VPI symptom has occurred, regardless of the target sound. For example, Idx 69.
ii. If the sound identified, see *Sound* column, has Idx value 14, 15, 16, 17,19 or 77 then a VPI symptom has occurred if the target sound is listed in the column *VPI Symptoms present if.*
iii. If criteria i or ii are not met, and the sound identified is listed anywhere in the column *VPI Symptoms present if* then*,* regardless of the target sound, a VPI symptom is present.

18.5.2.1.3 5 year mapping from SPEAK output

In the dataset (*Appendix Table 31)* the responses to Correct of "Yes" or "No" are provided in the column *WordAssessmentStatus* as "Correct" or "Incorrect" respectively. If "Incorrect" is provided then the sound that the child uttered is provided in the *ActualSounds* column and the SoundIdx provided in the column *ActualSoundsIdx*. If the *WordAssessmentStatus* column states "MediaNotAvailable" then this word was not available for assessment. If the *WordAssessmentStatus* column states "CouldNotAssess" or "MediaNotAvailable" these words should be removed from the child's assessment as outcomes cannot be derived from these assessments. The target sound for each word is provided in the 5 year output in the *TargetSounds* column with the IDX identified in the *TargetSoundsIdx* column.

18.5.2.1.4 5 year phonetics outcome derivation

PCC, PCP, PCM, oral consonant errors, non-oral consonant errors and VPI at 5 years of age are calculated as in *Section 18.5.2.1.2* using *Appendix Table 36* with column headings and content changed as described in *Section 18.5.2.1.3.*

18.5.2.1.5 Criteria for individual assessment to be included in outcome analysis

Phonetic recordings (Recording ID: 3y_1 and 5y_1, *Table 4*) require a child to attempt 30 and 36 pre-specified words (target sounds) at the 3 and 5 year timepoint respectively. Words

classified as “Could Not Assess” will not be included. A child must have attempted at least 50% (15 at 3 year, 18 at 5 year) words for inclusion in the outcome set.

At the March speech meeting SLTs provided feedback on some words they ‘Could not assess’ or struggled to assess due to the editing of the recording. The AC team and CSG evaluated the words the SLTs had reported, the CSG then decided if they should reassess the same recording, a recut recordings or no recut possible. The following should be applied:

- Where a word was reassessed and not recut the latest assessment of the word should be used in the outcome derivation. If an intra-rater/inter-rater assessment has been made this should be unaffected as the same material has been reassessed.
- Where a word was reassessed and recut the latest assessment of the word should be used in the outcome derivation. If an intra-rater assessment has been made then the ‘initial’ assessment for this word should be set to ‘Could not assess’ as the assessment of the different material can not be compared. If an inter-rater assessment has been made, ensure that both SLTs have assessed the same material.

Where two assessments have been made on the same child by the same assessor then the following criteria for selection applies:

- *Critieria one*: if both assessments meet the criteria of at least 50% of words asessed, then the latest assessment is used.
- *Criteria two*: if one assessment has less than 50% of words assessed and the other has at least 50% of words assessed, the assessment with at least 50% of words assessed is used.

Where two assessments have been made on the same child by different assessors then the following criteria for selection applies:

- *Criteria one*: if one assessment has less than 50% of words assessed and the other at least 50% of words assessed then use the one with at least 50% of words assessed.
- *Criteria two*: if both assessments meet the criteria of at least 50% of words assessed and both assessors are independent of the child i.e. not a therapist at the same centre that randomised the child, then an assessment will be randomly selected.
- *Criteria three*: if both assessments meet the criteria of at least 50% of words assessed and one of the two assessors is independent of the child i.e. not a therapist at the same centre that randomised the child, then the independent assessment will be used.

- *Criteria four*: if both assessments meet the criteria of at least 50% of words assessed and neither assessors are independent of the child i.e. not a therapist at the same centre that randomised the child, then an assessment will be randomly selected.

18.5.2.2 VPC-rate assessments

For VPC-rate assessment, an assessor can review if a child is either: Competent, Marginally Incompetent or Incompetent through reviewing spontaneous speech or bus story (bus story collected at 5 years only). Example output is provided *in Appendix Table 30* and *Appendix Table 33*. If an assessor deemed that they "Could Not Assess" the recording a reason was provided, free text will be categorised as advised by the CSG. Three assessors will assess VPC-rate at 3 and 5 years of age, the majority classification will be used as in

Table 15. Where the same assessor has made more than one assessment on a child, the latest assessment will be used.

Table 15: Majority classification of VPC-rate

Outcome classifcation (score)	Majority size	Number of SLTs per classification		
		Competent	Marginally incompetent	Incompetent
Competent (0)	3	3	0	0
	2	2	1	0
	2	2	0	1
Marginally competent (1)	3	0	3	0
	2	1	2	0
	2	0	2	1
	1	1	1	1
Incompetent (2)	3	0	0	3
	2	1	0	2
	2	0	1	2

The majority assessment will be used to determine whether a patient has insufficient VPC-rate ("Incompentent" according to the majority) or sufficient VPC-rate ("Marginally Incompetent" or "Competent" according to the majority).

18.5.2.2.1 3 year mapping from SPEAK output

In the output (*Appendix Table 30)* if the column *Step1* is "Yes" then the child is "Competent". However if the column states "No" then the column *Step2* will indicate if the child is "Marginally Incompetent" or "Incompetent". However, if the column *Step1* states "Could not assess" then a reason should be provided in the column *Comments* where a free text reason should be provided.

18.5.2.2.2 3 year VPC-rate outcomes derivation

Each of the three assessments per child will be as reported by the assessor classification. The child will then be classified according to the majority assessment.

18.5.2.2.3 5 year mapping from SPEAK output

In the output *(Appendix Table 33)* if the column *WithinNormalLimits* is "Yes" then the child is "Competent". However, if the column states "No" then the *DegreeOfSeverity* column will indicate if the child is "Marginally Incompetent" or "Incompetent". However, if the column *WithinNormalLimits* states "Could not assess" then a reason should be provided in the column *CouldNotAssessReason* where "other" is selected a free text reason should be provided in *OtherText* column.

18.5.2.2.4 5 year VPC-rate outcomes derivation

Each of the three assessments per child will be as reported by the assessor classification. The child will then be classified according to the majority assessment.

18.5.2.2.5 Number of assessments required

Two complete assessments of the outcome are needed to be included in the outcome set. Therefore two or more assessments of "Could Not Assess" will not be included. Where there is one assessment of "Could Not Assess" and two complete assessments then the following applies:

- If the two assessments agree then report as the outcome
- If the two assessments completely disagree (i.e. Competent and Incompetent) then classify Marginal Incompetent.
- If the two assessments partially agree (i.e. Competent and Marginal Incompetent or Marginal Incompetent and Incompetent) then report as the most severe (i.e. Marginal and Incompetent respectively).

18.5.2.3 Hypernasality assessment

SPEAK at 5 years was updated to allow for the assessment of hypernasality. Hypernasality is a single recording comprising the first nine of the total 36 target words concatenated for each child. These nine words have been selected by the CSG to be phonetically similar across the five languages. If a word is not attempted by the child the string will be shorter than nine words. The nine-word string must consist of at least 5 words to be included in the analysis. Each 9-word string will be assessed by three assessors, whereby each assessment contains a score of 0 to 3 depending on the assessor's judgement of "Within Normal Limits" and "Degree of Severity". Example output is provided in *Appendix Table 32.*

From the software output (*Appendix Table 32)* the column titled *WithinNormalLimits* will specify “Yes”, ”No” or “Could not assess”. Where “No” is selected the *DegreeOfSeverity* should state the hypernasality of each patient as either “Mild”, “Moderate” or “Severe”. If the *WithinNormalLimits* column specifies “Could not assess” then the reason should be stated in the column *CouldNotAssessReason*, if “Other” was selected this will be provided as free text in the column *OtherText*. The score derivation is indicated in *Table 16* and combined into a majority assessment as shown in *Table 17.*

Where two assessments have been made by the same assessor the latest assessment will be used.

Table 16: Hypernasality score

SPEAK output		Hypernasality score assigned
WithinNormalLimits	DegreeOfSeverity	
Yes	N/A	0
No	Mild	1
No	Moderate	2
No	Severe	3

Table 17: Classification of hypernasality using multiple assessments

Number of SLTs making assessment				Majority classification	
Within normal limits (0)	Not within normal limits, hypernasality:				
	Mild (1)	Moderate (2)	Severe (3)	Outcome	Score
3	0	0	0	Within Normal Limits	0
2	1	0	0		
2	0	1	0		
2	0	0	1		
1	2	0	0	Mild hypernasality	1
1	1	1	0		
1	1	0	1		
0	3	0	0		
0	2	1	0		
0	2	0	1		
1	0	2	0	Moderate hypernasality	2
1	0	1	1		
0	1	2	0		
0	1	1	1		
0	0	3	0		
0	0	2	1		
1	0	0	2	Severe hypernasality	3
0	1	0	2		
0	0	0	3		
0	0	1	2		

18.5.2.3.1 Number of assessments required

Two complete assessments of the outcome are needed to be included in the outcome set. Therefore two or more assessments of “Could Not Assess” will not be included. Where there is one assessment of “Could Not Assess” and two complete assessments then the following applies:

- If the two assessments agree then report as the outcome
- If two assessments disagree then the outcome is derived following *Table 18.*

Table 18 Two Hypernasality Assessments

Assessment 1	Assessment 2	Outcome
Within Normal Limits	Mild Hypernasality	Mild Hypernasality
Within Normal Limits	Moderate Hypernasality	Mild Hypernasality
Within Normal Limits	Severe Hypernasality	Moderate Hypernasality
Mild Hypernasality	Moderate Hypernasality	Moderate Hypernasality
Mild Hypernasality	Severe Hypernasality	Moderate Hypernasality
Moderate Hypernasality	Severe Hypernasality	Severe Hypernasality

18.6 Analysis of outcomes

18.6.1 Primary outcome: Velopharyngeal function at five years of age

The primary endpoint (*PO_5y; Appendix Table 23*) is insufficient velopharyngeal function at 5 years of age. Insufficient velopharyngeal function is determined by a score of four or more on the VPC-sum. Assessments are made by reviewing speech endpoint sources (*5y_1 and 5y_3; Table 4*), *Section 5* of the TOPS Protocol provides more details.

The VPC-sum score has a six-category ordinal scale (values of one to six). It is based on the four main speech characteristics of velopharyngeal impairment. Further VPC-sum scoring details are in *Section 9.1.4.5.1* of the TOPS Protocol. A patient is categorized using this score as follows:

- *Sufficient:* a score of 0, 1, 2 or 3 on the VPC-sum scale
- *Insufficient*: a score of 4, 5 or 6 on the VPC-sum scale.

18.6.1.1 Speech outcome assessment

Assessment completion and quality will be reported as in *Section 17.2* and *Section 18.4.4*.

18.6.1.2 Derivation

VPC-sum (*PO_5y; Appendix Table 23*) is comprised of three components: Hypernasality, Active non-oral errors *(SO_5d_5y; Appendix Table 23)* and VPI symptoms.

18.6.1.2.1 Deriving individual components for VPC-sum calculation

18.6.1.2.1.1 Hypernasality

Hypernasality is derived as described in *Section 18.5.2.3*. Hypernasality scores will be reported in accordance with the majority of three assessments, see *Table 17.*

The *score column* in *Table 17* gives the hypernasality assessment score to be converted to a component score as in *Figure 1.*

18.6.1.2.1.2 Active non-oral errors

This is derived as in *Section 18.5.2.1.3* and *18.5.2.1.4*. The total number of non-oral errors will be categorised and assigned a score as in *Figure 1.*

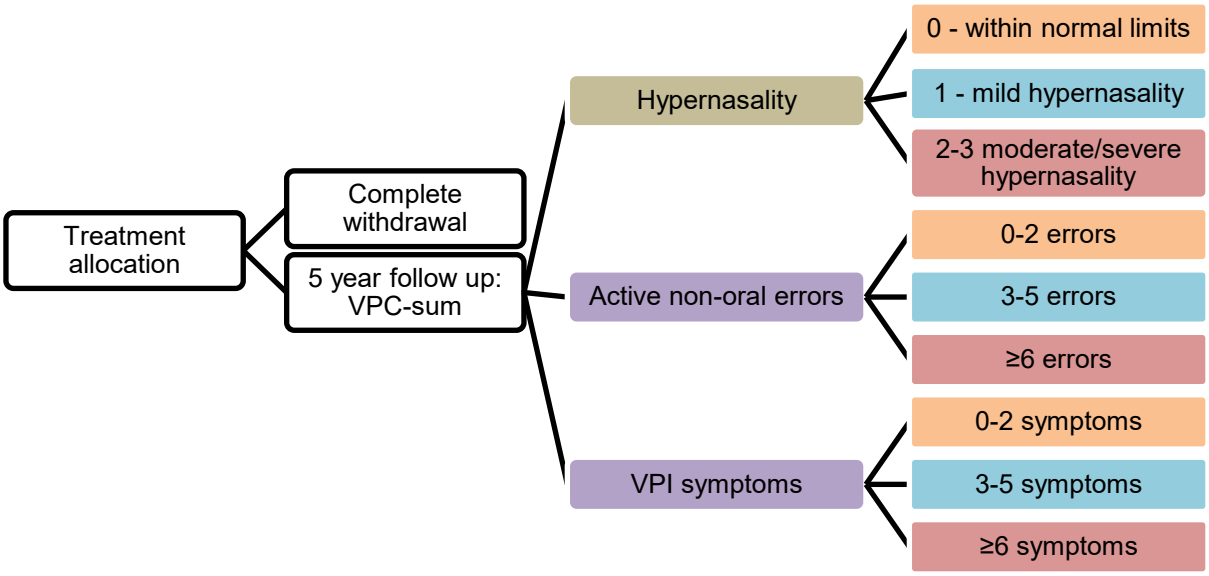
18.6.1.2.1.3 VPI-symptoms

VPI symptoms are derived as in *Section 18.5.2.1.3* and *18.5.2.1.4*. The total number of VPI symptoms will be categorised and assigned a score as in *Figure 1.*

18.6.1.2.2 Combining individual scores for VPC-sum calculation

Figure 1 and *Table 19* provides the transfer rules for each component (Hypernasality, Active non-oral errors and VPI symptoms) to create the VPC-sum, each component contributes a score of zero to two for the VPC-sum score which ranges from zero to six. The score of zero for each component indicates the following: Hypernasality within normal limits, zero to two active non-oral errors and zero to two VPI symptoms. *Figure 1* summarises the calculation of VPC-sum from each of the three components used.

Figure 1: VPC-sum component score



Recording Type	VPC-sum component score
9-word string	0
Phonetics	1
	2

Table 19: Transfer rules for VPC-sum

Variables	Transfer rules			VPC-sum component
A: Hypernasality	0 = **0**	1 = **1**	2-3 =**2**	**A [0-2]**
B: Active non-oral errors	0-2 =**0**	3-5 = **1**	≥6 = **2**	**B [0-2]**
C: VPI-symptoms	0-2 =**0**	3-5 = **1**	≥6 = **2**	**C [0-2]**

The VPC-sum is calculated by adding the components A, B and C: VPC-sum = A + B + C.

18.6.1.3 Analysis

18.6.1.3.1 Primary analysis

The following will be reported overall and split by treatment arm. Summaries will be reported as frequencies and percentages, defined as follows:

- Number of participants randomised (i)
- Number of participants with recordings eligible for assessment (ii, denominator i), reasons for missing data will be provided
    - Both phonetics and nine word string (iii, denominator i)
    - Phonetics only (iv, denominator i)

  - Nine word string only (v, denominator i)
  - Neither phonetics or nine word string (vi, denominator i)
- Number of participants with assessments that:
  - Were made using audio files only, see *Section 17.2.* (vii, denominator ii,iii iv,v)
  - Were outside the acceptable time window, see *Section 17* (viii, denominator ii, iii, iv,v)
  - For phonetics assessments only, contain greater than or equal to 18 words, see *Section 18.5.2.1.5* (ix, denominator ii,iii, iv)
  - For nine word string only, contain greater than five words (x, denominator ii,iii, v, vi)
- Number of participants included in primary outcome set, with reasons provided where not clear from categories above e.g. assessment not possible due to sound (xi, denominator ii)
- Number of participants with each subscore category (denominator xi):
  - Hypernasality (0, 1, 2)
  - Active non-oral errors (0, 1, 2)
  - VPI-symptoms (0, 1, 2)
- Number of participants with each overall score (0 - 6) (xii, denominator xi)

The number of participants who are classified as having (a) sufficient (score 0 to 3) and (b) insufficient (score 4 to 6) VPC-sum (and percentages) will be presented overall and for each treatment arm and primary language of recruiting site (Brazil, Denmark, Norway, Sweden, UK). Summaries will use the denominators of the number included in the outcome set (xi), the number randomised to the treatment arm and the number from each primary language in the primary outcome set.

The relative risk[34] together with 95% confidence interval (95% CI) will be reported along with a two-sided p-value from a chi-square test. **This will be the formal comparison of the primary outcome as defined in the TOPS Protocol.**

As an exploratory analysis, a cross table will present (a) and (b) split by the number of participants requiring secondary surgery due to velopharyngeal insufficiency (as recorded on *Form 12: Secondary surgery*). Summaries will use the denominators of the number included in the outcome set (xi) and the number randomised to the treatment arm in the primary outcome set. The number in each treatment group that have (c) sufficient (score 0 to 3) VPC-sum and have had no secondary surgery due to velopharyngeal insufficiency will be compared

with (d) insufficient (score 4 to 6) VPC-sum or have had a secondary surgery due to velopharyngeal insufficiency. The relative risk[34] together with 95% confidence interval (95% CI) will be reported along with a two-sided p-value from a chi-square test. **This will not constitute a formal outcome comparison because the definition of the groups compared is not compatable with any outcome defined in the TOPS Protocol.**

18.6.1.3.2 Goodness of fit

In the circumstance that the sufficient and insufficient VPC-sum for each group contains less than five participants, thereby raising concern over the appropriateness of a chi-square test, then a Fisher's exact test will be used.

18.6.1.3.3 Sensitivity analysis

A sensitivity analysis will be performed to check the robustness of the results to the inclusion/exclusion set for speech recordings such that:

- Phonetic assessments containing less than 18 words are included;
- Hypernasality assessments containing less than 5 words are included;
- Phonetic assessments based on audio recordings only are excluded;
- Any assessments made on recordings taken before/after the follow up window will be excluded.

As a result of blind review, two additional sensitivity analyses will be performed, instead of a per protocol analysis, to test the robustness of the results to the two most important protocol deviations:

- Patients with protocol deviation 3, Infants with syndromic cleft palate or severe developmental delay - Identification post randomisation, will be removed from the primary outcome set
- Patients with protocol deviation 7, surgery outside the expected time window, will be removed from the primary outcome set

18.6.1.3.4 Heterogeneity by surgeon

Heterogeneity in treatment between surgeons will be accounted for using a multilevel logistic regression model (for insufficient VPC-sum) including; surgeon who performed the surgery, minimisation factor size of cleft at baseline, treatment allocation, and an intercept.

18.6.1.3.5 Inter-rater reliability

VPC-sum is a component score deriving from the assessment of two different recording types: nine-word string and phonetics, see *Figure 1*. The nine-word string has three independent assessors, see *Section 18.5.2.3.* Phonetics has one assessor for all with approximately 50% independently assessed by a second assessor. As the raters are not the same across each recording, exploring inter-rater agreement on the overall score is not possible. Instead, inter-rater reliability will be explored separately for each subcomponent basis: hypernasality, active non-oral errors, and VPI-symptoms.

18.6.1.3.5.1 Hypernasality

Hypernasality as a subcomponent is a categorical outcome. Inter-rater analysis will be performed by assessing agreement between assessments of each classification ("0: within normal limits", "1: mild hypernasality", and "2-3: moderate/severe hypernasality", see *Figure 1*). The level of agreement between the three assessors will be summarised as frequencies and percentages (denominator; the number with two Hypernasality assessments) for each of the possible combinations of agreement, see *Table 17.*

18.6.1.3.5.2 Active non-oral errors

Active non-oral errors as a subcomponent is a categorical outcome. Inter-rater analysis will be performed by assessing agreement between assessments of each classification only on the subset of recordings with two assessors. Analysis will explore agreement between assessments of each classification ("0-2 errors", "3-5 errors", and "≥6 errors", see *Figure 1*). The level of agreement will be summarised as frequencies and percentages (denominator; the number with two Phonetics assessments) for each of the possible combinations of agreement.

18.6.1.3.5.3 VPI-symptoms

VPI-symptoms as a subcomponent is a categorical outcome ("0-2 symptoms", "3-5 symptoms", and "≥6 symptoms", see *Figure 1*). Agreement will be summarised using the methods described in *Section 18.6.1.3.5.2.*

18.6.1.3.6 Intra-rater reliability

VPC-sum is a component score deriving from the assessment of two different recording types: nine-word string and phonetics, see *Figure 1*. Approximately 25% of the nine-word string recordings will be reassessed by the same assessor. Approximately 20% of the phonetics recordings will be reassessed by the same assessor. As the raters are not the same across

each recording, and the reassessment approach is dependent on the recording type, exploring intra-rater agreement on the overall score is not possible. Instead, intra- rater reliability will be explored separately for each subcomponent basis: hypernasality, active non-oral errors, and VPI-symptoms.

18.6.1.3.6.1 Hypernasality

Hypernasality as a subcomponent is a categorical outcome. Intra-rater analysis will be performed by assessing agreement between assessments of each classification only on the subset of recordings with two assessments.

Intra-rater analysis will be performed by assessing agreement of each classification ("0: within normal limits", "1: mild hypernasality", and "2-3: moderate/severe hypernasality", see *Figure 1*) between the first and second assessment. Frequencies and percentages (denominator; the number with a repeat assessment) will be presented using a cross table.

18.6.1.3.6.2 Active non-oral errors

Agreement will be summarised using the methods described in *Section 18.6.1.3.6.1* for categories: "0-2 errors", "3-5 errors", and "≥6 errors", see *Figure 1.*

18.6.1.3.6.3 VPI-symptoms

Agreement will be summarised using the methods described in *Section 18.6.1.3.6.1* for categories: "0-2 symptoms", "3-5 symptoms", and "≥6 symptoms", see *Figure 1.*

18.6.2 Secondary outcomes

18.6.2.1 Secondary outcome 1: Velopharyngeal function at age 5 years

This outcome is measured using the following endpoints:

a) Velopharyngeal composite score summary (VPC-sum)
b) Insufficient velopharyngeal function (VPC-rate)

18.6.2.1.1 Derivation

a) Velopharyngeal composite score summary (VPC-sum)

VPC-sum *(Outcome ID: SO_1a_5y, Appendix Table 23)* uses the individual score that contributes to the binary classification for the primary outcome (*Outcome ID: PO_5y, Appendix Table 23*). This outcome is therefore derived as in *Section 18.6.1.*

b) Insufficient velopharyngeal function (VPC-rate)

VPC-rate (*Outcome ID: SO_1b_5y, Appendix Table 23)* is derived as in Section *18.5.2.2.*

18.6.2.1.2 Analysis

a) Velopharyngeal composite score summary (VPC-sum)

VPC-sum will be summarised overall and for each treatment arm and primary language of recruiting site (Brazil, Denmark, Norway, Sweden, UK), by means and standard deviations (SDs), medians and inter-quartile range (IQR) if data are skewed. Minimum and maximum values will also be presented.

If the data are normally distributed (assessed using QQ plot by treatment group), the mean VPC-sum will be tested using a t-test, and the mean difference and 95% CI reported. Otherwise, the mean score will be tested using a non-parametric equivalent test.

Update following blind review:
Data is skewed for the outcome therefore, VPC-sum will be summarised by medians and IQR and analysed using a Mann Whitney U test.

b) Insufficient velophayngeal function (VPC-rate)

The following will be reported overall and split by treatment arm. Summaries will be reported as frequencies and percentages, defined as follows:

- Number of participants randomised (i)
- Number of participants with VPC recordings eligible for assessment, reasons for missing data will be provided (ii, denominator i)
- Number of participants with assessments that:

    - Were made using audio files only, see *Section 17.2.* (iii, denominator ii)
    - Were outside the acceptable time window, see *Section 17* (iv, denominator, ii)
    - At least two SLT classified as 'Could not assess', where available further detail will be provided (v, denominator ii)
- Number of participants included in secondary outcome set, with reasons provided where not clear from categories above e.g. assessment not possible due to sound (vi, denominator i and ii)
- Number of participants assessed using:
    - Bus story (vii, denominator vi)
    - Spontaneous Speech (viii, denominator vi)
    - Combination of bus story/spontaneous speech (ix, denominator vi)
- Number of participants classified (according to the majority) as:
    - Incompetent (x, denominator vi)
    - Marginally incompetent (xi, denominator vi)
    - Competent (xii, denominator vi)

The number of participants (and percentages) who are classified as having VPC-rate of insufficient ("Incompentent", see *Table 15*) and (b) sufficient ("Marginally Incompetent", "Competent", see *Table 15*

Table 15) will be presented overall and for each treatment arm and primary language of recruiting site (Brazil, Denmark, Norway, Sweden, UK). Using the denominators of the number included in the outcome set (vi), the number randomised to the treatment arm and the number from each primary language in the secondary outcome set.

The relative risk[34] together with 95% confidence interval (95% CI) will be reported along with a two-sided p-value from a chi-square test.

In the circumstance that the sufficient and insufficient VPC-rate for each group contains less than five participants, thereby raising concern over the appropriateness of a chi-square test, a Fisher's exact test will be used.

No sensitivity analysis will be performed.

For exploratory purposes, VPC-rate on the three point scale ("Incompentent" "Marginally Incompetent", "Competent") will be compared between groups using a chi-squared test for trend. A two-sided p-value will be reported. In the circumstance that the numbers within each

categorisation for each group contains less than five participants, thereby raising concern over the appropriateness of a chi-square test for trend, a proportional odds model will be applied. **Results of this test is intended to be exploratory and will not replace the main outcome measure of VPC-rate on a dichotomous scale.**

18.6.2.1.2.1 Inter-rater reliability

a) Velopharyngeal composite score summary (VPC-sum)

As VPC-sum uses the individual score that contributes to the binary classification for the primary outcome (*Outcome ID: PO_5y, Appendix Table 23*), inter-rater reliability for this outcome will be performed as part of the primary outcome assessment, see *Section 18.6.1.3.5*.

b) Insufficient velophayngeal function (VPC-rate)

Inter-rater analysis will be performed by assessing agreement between assessments of each classification (insufficient or sufficient, see *Section 18.6.2.1.2*). The level of agreement between the three assessors will be summarised as frequencies and percentages (denominator; the number of recordings with two assessments) for each of the possible combinations of agreement, see Table 15.

18.6.2.1.2.2 Intra-rater reliability

a) Velopharyngeal composite score summary (VPC-sum)

As VPC-sum *(Outcome ID: SO_1a_5y, Appendix Table 23)* uses the individual score that contributes to the binary classification for the primary outcome (*Outcome ID: PO_5y, Appendix Table 23*), intra-rater reliability for this outcome will be performed as part of the primary outcome assessment, see *Section 18.6.1.3.6*.

b) Insufficient velophayngeal function (VPC-rate)

Approximately 25% of the VPC-rate recordings will be reassessed by the same assessor. Intra-rater analysis will be performed by assessing agreement between assessments of each classification only on the subset of recordings with two assessments. Intra-rater analysis will be performed by assessing agreement of each classification (insufficient or sufficient, see *Section 18.6.2.1.2*) between the first and second assessment. Frequencies and percentages (denominator; the number of repeat assessments) will be presented using a cross table.

18.6.2.2 Secondary outcome 2: Velopharyngeal function at 3 years

This outcome is measured using the following end points:

a) Insufficient velopharyngeal function (VPC-rate)

b) Velopharyngeal insufficiency symptoms

18.6.2.2.1 Derivation

a) Insufficient velopharyngeal function (VPC-rate)

VPC-rate (*Outcome ID: SO_2a_3y, Appendix Table 23*) is derived as in *Section 18.5.2.2.*

b) Velopharyngeal insufficiency symptoms

Velopharyngeal insufficiency symptoms (VPI symptoms, *Outcome ID: SO_2b_3y, Appendix Table 23*) is derived as in *Section 18.5.2.1*. Velopharyngeal insufficiency symptoms is calculated as follows:

$$VPI\ symptoms = \frac{number\ of\ VPI\ symptoms}{number\ of\ words\ attempted}$$

18.6.2.2.2 Analysis

a) Insufficient velopharyngeal function (VPC-rate)

VPC-rate (*Outcome ID: SO_2a_3y, Appendix Table 23*) is analysed as in *Section 18.6.2.1.2.*

b) Velopharyngeal insufficiency symptoms

The following will be reported overall and split by treatment arm. Summaries will be reported as frequencies and percentages, defined as follows:

- Number of participants randomised (i)
- Number of participants with recordings eligible for assessment, reasons for missing data will be provided
    - Phonetics recording (ii, denominator i)
    - No phonetics recording (iii, denominator i)
- Number of participants with assessments that:
    - Were made using audio files only, see *Section 17.2.* (iv, denominator ii)
    - Were outside the acceptable time window, see *Section 17* (v, denominator ii)
    - For phonetics assessments only, contain at least 15 words, see *Section 18.5.2.1.5* (vi, denominator ii)
- Number of participants included in secondary outcome set (vii, denominator ii), with reasons provided where not clear from categories above e.g. assessment not possible due to sound

VPI symptoms will be summarised overall and for each treatment arm and primary language of recruiting site (Brazil, Denmark, Norway, Sweden, UK), by means and standard deviations (SDs), medians and inter-quartile range (IQR) if data are skewed. Minimum and maximum values will also be presented.

If the data is normally distributed (assessed using QQ plot by treatment group), the mean VPI symptoms will be tested using a t-test, and the mean difference and 95% CI reported. Otherwise, the mean score will be tested using a non-parametric equivalent test.

No sensitivity analysis will be performed.

Update following blind review:
Data is skewed for the outcome therefore, VPI symptoms will be summarised by medians and IQRand analysed using a Mann Whitney U test.

18.6.2.2.2.1 Inter-rater reliability

a) Insufficient velopharyngeal function (VPC-rate)

Inter-rater reliability of VPC-rate will be performed as in *Section 18.6.2.1.2.1.*

b) Velopharyngeal insufficiency symptoms

VPI symptoms are a continuous outcome (in the form of a percentage). Inter-rater analysis will be performed by assessing agreement between assessments of each classification only on the subset of recordings with two assessors (approximately 50%).

Agreement will be defined as the same outcome for an individual word. For example if two different sounds were identified by each assessor, but both resulted in a VPI symptom, then this would classify as an agreement for that word. Likewise, if different sounds are identified but if one does and one does not result in a VPI symptom then this would be a disagreement. The total number of VPI symptoms and percentage of words that assessors agree on, per child, will be summarised as means and standard deviations (SDs), or medians and inter-quartile range (IQR) if data are skewed. Minimum and maximum values will also be presented. If a word is classified as ‘Could not assess’ by one assessment and not by the other this word should not be compared for inter-rater reliability as the outcome can not be derived from the assessment.

18.6.2.2.2.2 Intra-rater reliability

a) Insufficient velopharyngeal function (VPC-rate)

Intra-rater reliability of VPC-rate will be performed as in *Section 18.6.2.1.2.2.*

b) Velopharyngeal insufficiency symptoms

VPI symptoms are a continuous outcome (in the form of a percentage). Intra-rater analysis will be performed by assessing agreement between assessments of each classification only on the subset of recordings with two assessments (approximately 20%).

Agreement is defined as in *Section 18.6.2.2.2.1.*

18.6.2.3 Secondary outcome 3: Canonical babbling at age 12 months

This outcome is measured using the following end points:

a) Canonical babbling present
b) Canonical babbling ratio
c) Consonant inventory

18.6.2.3.1 Derivation

a) Canonical babbling present

Canonical babbling present (*Outcome ID: SO_3a_12m, Appendix Table 23*) is derived as in *Section 18.5.1.1.*

b) Canonical babbling ratio

Canonical babbling ratio (*Outcome ID: SO_3b_12m, Appendix Table 23*) is derived as in *Section 18.5.1.2.*

c) Consonant inventory

Consonant inventory (*Outcome ID: SO_3b_12m, Appendix Table 23*) is derived as in *Section 18.5.1.3.*

18.6.2.3.2 Analysis

a) Canonical babbling present

The following will be reported overall and split by treatment arm. Summaries will be reported as frequencies and percentages, defined as follows:

- Number of participants randomised (i)
- Number of participants with spontaneous speech (*Recording ID: 12m_1*, *Table 4*) recordings eligible for assessment (ii, denominator i), reasons for missing data will be provided
- Number of participants with assessments that:
    - Were made using audio files only, see *Section 17.2.* (iii, denominator ii)
    - Were outside the acceptable time window, see *Section 17* (iv, denominator ii)
- Number of participants included in secondary outcome set, with reasons provided where not clear from categories above (vi, denominator ii)
- Number of participants classified (according to the majority) as:
    - Canonical (vii, denominator vi)
    - Not canonical (viii, denominator vi)

The number of participants (and percentages) who are classified as being “Canonical” or “Not canonical“ (see *Table 12*) will be presented overall and for each treatment arm and primary language of recruiting site (Brazil, Denmark, Norway, Sweden, UK). Using the denominators of the number included in the outcome set (vi), the number randomised to the treatment arm and the number from each primary language in the secondary outcome set.

The relative risk[34] together with 95% confidence interval (95% CI) will be reported along with a two-sided p-value from a chi-square test.

In the circumstance that the “Canonical” or “Not canonical“ categorisations for each group contains less than five participants, thereby raising concern over the appropriateness of a chi-square test, a Fisher’s exact test will be used.

No sensitivity analysis will be performed.

An exploratory analysis, comparing classification of canonical babbling present according to each approach for classification, the main outcome, see *Section 18.5.1.1* and the alternative, see *Section 18.5.1.1.1,* will be compared using frequencies and percentages in a cross table.

b) Canonical babbling ratio

The canonical babbling present summary, see *Section 18.6.2.3.2*, will present the cohort of eligible participants for this outcome.

Canonical babbling ratio will then be summarised, presented overall and for each treatment arm and primary language of recruiting site (Brazil, Denmark, Norway, Sweden, UK), by means and standard deviations (SDs), medians and inter-quartile range (IQR) if data are skewed. Minimum and maximum values will also be presented.

If the data are normally distributed (assessed using QQ plot by treatment group), the mean consonant inventory size will be tested using a t-test, and the mean difference and 95% CI reported. Otherwise, the mean score will be tested using a non-parametric equivalent test.

No sensitivity analysis will be performed.

Update following blind review:

Data is normally distributed for the outcome therefore, canonical babbling ratio will be summarised by means and SDs. The mean score will be tested using a t-test.

c) Consonant inventory

Consonant inventory will be analysed in the same way as canonical babbling ratio, see *Section 18.6.2.3.2.*

Update following blind review:

Data is normally distributed for the outcome therefore, consonant inventory size will be summarised by means and SDs. The mean score will be tested using a t-test.

18.6.2.3.2.1 Inter-rater reliability

a) Canonical babbling present

Inter-rater analysis will be performed by assessing agreement between assessments of each classification ("Canonical", "Not canonical"). The level of agreement between the three assessors will be summarised as frequencies and percentages (denominator vi) for each of the possible combinations of agreement, see *Table 12*.

b) Canonical babbling ratio

Inter-rater analysis will be performed by calculating the maximum difference between the three assessments. This will be calculated as the largest CBR minus the minimum CBR assessed for each participant. The maximum difference will be summarised by means and standard deviations (SDs), medians and inter-quartile range (IQR) if data are skewed. Minimum and maximum values will also be presented.

c) Consonant inventory

Consonant inventory is a continuous outcome, yet the outcome relies on the identification of unique syllables. Inter-rater analysis will therefore be performed by presenting all unique syllables against the number of children that uttered the syllables as identified by none, one, two, three, at least one and at least two. Frequencies and percentages will be presented with the total number of children as the denominator.

18.6.2.3.2.2 Intra-rater reliability

Due to concerns of time required to conduct assessments, there were no formal intra assessments performed on the 12 month recordings. Therefore, no intra analysis will be undertaken.

18.6.2.4 Secondary outcome 4: Articulation at 3 years

This outcome is measured using the following end points:

a) Percent consonants correct (PCC)
b) Percent correct placement (PCP)
c) Percent correct manner (PCM)
d) Non-oral consonant errors
e) Oral consonant errors

18.6.2.4.1 Derivation

a) Percent consonants correct (PCC)

Percent consonants correct (*Outcome ID: SO_4a_3y, Appendix Table 23*) is derived as in *Section 18.5.2.1.2* and in accordance to the equation in *Table 20.*

b) Percent correct placement (PCP)

Percent correct placement (*Outcome ID: SO_4b_3y, Appendix Table 23*) is derived as in *Section 18.5.2.1.2* and in accordance to the equation in *Table 20.*

c) Percent correct manner (PCM)

Percent correct manner (*Outcome ID: SO_4c_3y, Appendix Table 23*) is derived as in *Section 18.5.2.1.2* and in accordance to the equation in *Table 20.*

d) Non-oral consonant errors

Non-oral consonant errors (*Outcome ID: SO_4d_3y, Appendix Table 23*) is derived as in *Section 18.5.2.1.2* and in accordance to the equation in *Table 20.*

e) Oral consonant errors

Oral consonant errors (*Outcome ID: SO_4e_3y, Appendix Table 23*) is derived as in *Section 18.5.2.1.2* and in accordance to the equation in *Table 20.*

Table 20: Equations to calculate articulation outcomes

Outcome	Derivation	Mapping structure column
PCC	$\frac{n\ sounds\ correct}{n\ sounds\ attempted}$x100	Identified by assessing SLT
PCP	$\frac{n\ sounds\ correct\ place}{n\ sounds\ attempted}$x100	Place of articulation
PCM	$\frac{n\ sounds\ correct\ manner}{n\ sounds\ attempted}$x100	Manner of articulation
Non-oral consonant errors	$\frac{n\ non\ oral\ errors}{n\ sounds\ attempted}$x100	Non-oral errors if
Oral consonant errors	$\frac{n\ oral\ errors}{n\ sounds\ attempted}$x100	Oral errors if

18.6.2.4.2 Analysis

a) Percent consonants correct (PCC)

Percent consonant correct will be analysed in the same way as VPI symptoms, see *Section 18.6.2.2.2.*

Update following blind review:

Data is skewed for the outcome therefore, PCC will be summarised by medians and IQR and analysed using a Mann Whitney U test.

b) Percent correct placement (PCP)

Percent correct placement will be analysed in the same way as VPI symptoms, see *Section 18.6.2.2.2.*

Update following blind review:

Data is skewed for the outcome therefore, PCP will be summarised by medians and IQR and analysed using a Mann Whitney U test.

c) Percent correct manner (PCM)

Percent correct manner will be analysed in the same way as VPI symptoms, see *Section 18.6.2.2.2.*

Update following blind review:

Data is skewed for the outcome therefore, PCM will be summarised by medians and IQR and analysed using a Mann Whitney U test.

d) Non-oral consonant errors

Non-oral consonant errors will be analysed in the same way as VPI symptoms, see *Section 18.6.2.2.2.*

Update following blind review:

Data is skewed for the outcome therefore, non-oral consonant errors will be summarised by medians and IQR and analysed using a Mann Whitney U test.

e) Oral consonant errors

Oral consonant errors will be analysed in the same way as VPI symptoms, see *Section 18.6.2.2.2.*

Update following blind review:

Data is skewed for the outcome therefore, Oral consonant errors will be summarised by medians and IQR and analysed using a Mann Whitney U test.

18.6.2.4.2.1 Inter-rater reliability

a) Percent consonants correct (PCC)

Inter-rater analysis of percent consonant correct will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.1.*

b) Percent correct placement (PCP)

Inter-rater analysis of percent correct placement will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.1.*

c) Percent correct manner (PCM)

Inter-rater analysis of percent correct manner will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.1.*

d) Non-oral consonant errors

Inter-rater analysis of non-oral consonant errors will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.1.*

e) Oral consonant errors

Inter-rater analysis of oral consonant errors will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.1.*

18.6.2.4.2.2 Intra-rater reliability

a) Percent consonants correct (PCC)

Intra-rater analysis of percent consonant correct will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.2.*

b) Percent correct placement (PCP)

Intra-rater analysis of percent correct placement will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.2.*

c) Percent correct manner (PCM)

Intra-rater analysis of percent correct manner will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.2.*

d) Non-oral consonant errors

Intra-rater analysis of non-oral consonant errors will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.2.*

e) Oral consonant errors

Intra-rater analysis of oral consonant errors will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.2.*

18.6.2.5 Secondary outcome 5: Articulation at 5 years

This outcome is measured using the following end points:

a) Percent consonants correct (PCC)
b) Percent correct placement (PCP)
c) Percent correct manner (PCM)
d) Non-oral consonant errors
e) Oral consonant errors

18.6.2.5.1 Derivation

a) Percent consonants correct (PCC)

Percent consonants correct (*Outcome ID: SO_5a_5y, Appendix Table 23*) is derived as in *Section 18.5.2.1.4* and in accordance to the equation in *Table 20.*

b) Percent correct placement (PCP)

Percent correct placement (*Outcome ID: SO_5b_5y, Appendix Table 23*) is derived as in *Section 18.5.2.1.4* and in accordance to the equation in *Table 20.*

c) Percent correct manner (PCM)

Percent correct manner (*Outcome ID: SO_5c_5y, Appendix Table 23*) is derived as in *Section 18.5.2.1.4* and in accordance to the equation in *Table 20.*

d) Non-oral consonant errors

Non-oral consonant errors (*Outcome ID: SO_5d_5y, Appendix Table 23*) is derived as in *Section 18.5.2.1.4* and in accordance to the equation in *Table 20.*

e) Oral consonant errors

Oral consonant errors (*Outcome ID: SO_5e_5y, Appendix Table 23*) is derived as in *Section 18.5.2.1.4* and in accordance to the equation in *Table 20.*

18.6.2.5.2 Analysis

a) Percent consonants correct (PCC)

Percent consonant correct will be analysed in the same way as VPI symptoms, see *Section 18.6.2.2.2*.

Update following blind review:

Data is skewed for the outcome therefore, PCC will be summarised by medians and IQR and analysed using a Mann Whitney U test.

b) Percent correct placement (PCP)

Percent correct placement will be analysed in the same way as VPI symptoms, see *Section 18.6.2.2.2*.

Update following blind review:

Data is skewed for the outcome therefore, PCP will be summarised by medians and IQR and analysed using a Mann Whitney U test.

c) Percent correct manner (PCM)

Percent correct manner will be analysed in the same way as VPI symptoms, see *Section 18.6.2.2.2*.

Update following blind review:

Data is skewed for the outcome therefore, PCM will be summarised by medians and IQR and analysed using a Mann Whitney U test.

d) Non-oral consonant errors

Non-oral consonant errors will be analysed in the same way as VPI symptoms, see *Section 18.6.2.2.2*.

Update following blind review:

Data is skewed for the outcome therefore, non-oral consonant errors will be summarised by medians and IQR and analysed using a Mann Whitney U test.

e) Oral consonant errors

Oral Consonant errors will be analysed in the same way as VPI symptoms, see *Section 18.6.2.2.2*.

Update following blind review:

Data is skewed for the outcome therefore, oral consonant errors will be summarised by medians and IQR and analysed using a Mann Whitney U test.

18.6.2.5.2.1 Inter-rater reliability

a) Percent consonants correct (PCC)

Inter-rater analysis of percent consonant correct will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.1.*

b) Percent correct placement (PCP)

Inter-rater analysis of percent correct placement will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.1.*

c) Percent correct manner (PCM)

Inter-rater analysis of percent correct manner will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.1.*

d) Non-oral consonant errors

Inter-rater analysis of non-oral consonant errors will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.1.*

e) Oral consonant errors

Inter-rater analysis of oral consonant errors will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.1.*

18.6.2.5.2.2 Intra-rater reliability

a) Percent consonants correct (PCC)

Intra-rater analysis of percent consonant correct will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.2.*

b) Percent correct placement (PCP)

Intra-rater analysis of percent correct placement will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.2.*

c) Percent correct manner (PCM)

Intra-rater analysis of percent correct manner will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.2.*

d) Non-oral consonant errors

Intra-rater analysis of non-oral consonant errors will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.2.*

e) Oral consonant errors

Intra-rater analysis of oral consonant errors will be assessed in the same way as VPI symptoms, see *Section 18.6.2.2.2.2.*

18.6.2.6 Secondary outcome 6: Post-operative/long term complications

This outcome is measured using the following end points:

a) Dehiscence
b) Infection
c) Fistula details (evidence, length, location, maximum width)

Other long term and post-operative complications will be reported as safety events, see *Section 21.*

18.6.2.6.1 Derivation

a) Dehiscence

Dehiscence (*Outcome ID: SO_6a_48h, SO_6a_30d, Appendix Table 23*), is collected at 48 hours and 30 days post surgery on the *Form 2: Surgery and Postoperative Care* and *Form 3: Late complications* CRF respectively. All events on these forms will count as dehicence events.

Dehisence is also collected on *Form 12: Secondary surgery*. Dehicence events on these forms associated with a secondary surgery will also count as an event.

A response of “Yes” to *Dehiscence* on these forms will indicate participants who reach this endpoint.

Any free text provided for *Other complication* on *Form 2 and Form 3* and free text provided on *Form 12* will be reviewed by the CInv, or delegated other, in order to identify any additional participants with this endpoint.

b) Infection

Infection *(Outcome ID: SO_6b_48h, SO_6b_30d, Appendix Table 23)*, is collected at 30 days post surgery on the *Form 3: Late complications* CRF. All events on these forms will count as infection events.

A response of “Yes” to *Infection* will indicate participants who reach this endpoint.

Any free text provided for *Other complication* on *Form 2 and Form 3* and free text provided on *Form 12* will be reviewed by the CInv, or delegated other, in order to identify any additional participants with this endpoint. Free text on *Form 2* will identify infection at 48 hours post surgery.

c) Fistula

Fistula *(Outcome ID: SO_6c_30d, SO_6c_3y, SO_6c_5y, Appendix Table 23)* is collected at 30 days post surgery and at the three and five years on the CRFs *Form 3: Late complications* and *Form 7: Fistula Assessment.* All events on these forms will count as fistula events. Fistula is also collected on *Form 12: Secondary surgery*. The date of secondary surgery will be used to classify the event into the next time point for fistula assessment: either 12 months, 3 years or 5 years. For example, if secondary surgery occurred at 18 months of age this event would be classified as being observed at the 3 year time point.

A response of "Yes" or "Probably" to *Evidence of fistula* will indicate participants who reach this endpoint.

Any free text provided for *Other complication* on *Form 2 and Form 3* and free text provided on *Form 12* will be reviewed by the CInv, or delegated other, in order to identify any additional participants with this endpoint.

Where there is evidence of fistula, length (mm), width (mm) and location of the fistula (soft palate / at border of soft and hard palate / hard palate) is also reported.

18.6.2.6.2 Analysis

a) Dehiscence

The following will be reported overall and split by treatment arm. Summaries will be reported as frequencies and percentages, defined as follows:

- Number of participants randomised (i)
- Number of participants with dehicence assessment (ii, denominator i), by group and timepoint
- Number of participants included in secondary outcome set (iii, denominator i), with reasons provided where not clear from categories above
- Number of participants classified as having:
    - Dehiscence (iv, denominator iii)
    - No dehiscence (v, denominator iii)

The number of participants (and percentages) who are classified as having "Dehiscence" or "No dehiscence" will be presented overall, and split by time point and treatment arm. Using the denominators of the number included in the outcome set (iii), the number randomised to the treatment arm in the secondary outcome set.

The relative risk[34] together with 95% confidence interval (95% CI) will be reported along with a two-sided p-value from a chi-square test.

In the circumstance that the “Dehiscence” or “No dehiscence“ categorisations for each group contains less than five participants, thereby raising concern over the appropriateness of a relative risk chi-square test, a Fisher’s exact test will be used.

No sensitivity analysis will be performed.

Histograms for each randomised group and each safety set will be presented for each follow up timepoint to present the distribution of events by analysis set.

b) Infection

Infection is analysed using the same approach as dehiscence, see *Section 18.6.2.6.2.*

c) Fistula

Fistula is analysed using the same approach as dehiscence, see *Section 18.6.2.6.2.* Fistula events were collected at 12 months for those who underwent surgery at 6 months, these will also be summarised.

Where there is evidence of fistula, descriptives of the fistula will be summarised overall, and split by timepoint and treatment group. The length (mm) and width (mm) will be summarised by means and standard deviations (SDs), medians and inter-quartile range (IQR) if data are skewed. Minimum and maximum values will also be presented. The location of the fistula (soft palate / at border of soft and hard palate / hard palate) will be summarised by frequencies and percentages (denominator; the number with a fistula assessed).

18.6.2.7 Secondary outcome 7: Hearing level

This outcome is measured using the following end points:

a) 12 months:
   i. Abnormal transient optoacoustic emission (TEOAE)
   ii. Abnormal Sound field audiometry
b) 3 and 5 years
   i. Abnormal pure tone audiometry in at least one ear
   ii. Abnormal pure tone audiometry in both ears
   iii. Severity of better ear (normal, mild, moderate, severe, profound)

18.6.2.7.1 Derivation

a) 12 months

Hearing level at 12 months is collected on *Form 6: 12 months audiology data collection.*

i. Abnormal transient optoacoustic emission (TEOAE)

Acoustic emission is tested at 12 months of age using the stand alone test TEOAE. This test has the following responses: “Yes”, “No” or “Not tested”. Reasons for “Not tested” should be provided as freetext and will be classified by the TOPS Lead Audiologist. If a response is missing it will be assumed the child is not tested, inconclusive or cannot be interpreted (agreed in meeting with TOPS Lead Audiologist on 22/05/2017).

Abnormal TEOAE *(Outcome ID: SO_7ai_12m, Appendix Table 23)* is defined as a response of “No” to the normal acoustic emission test.

ii. Abnormal Sound field audiometry

Sound field audiometry is tested at 500, 1000, 2000 and 4000 Hz. The response to each frequency is recorded in decibels hearing level (dB HL). Reasons for “Not tested” should be provided as freetext and will be classified by the TOPS Lead Audiologist.

Visual reinforcement audiometry (VRA) is tested at the same frequencies, at the same visit. VRA results will be used and interpreted in the same way as Sound field audiometry for participants who have not had Sound field audiometry assessed.

Update from blind review:
VRA should be used in the outcome derivation where provided and Sound field audiometry only used where VRA is not provided.

A minimum of one dB HL result at one frequency is required for inclusion in outcome comparison.

At any frequency, a threshold of 30 dB HL will determine if the patient has abnormal hearing levels (agreed in meeting with TOPS Lead Audiologist on 01/07/2019) such that:

- Values >30 dB HL will be categorised as “abnormal”;
- Values ≤30 dB HL will be categorised as “normal”.

A patient will be classified as “abnormal” Sound field audiometry at 12 months (*Outcome ID:* SO_7aii_12m*, Appendix Table 23*) if at least one frequency indicates this.

b) 3 and 5 years

Hearing level at 3 and 5 years is collected on *Form 8: Audiology data collection 3 and 5 years.*

i. Abnormal pure tone audiometry in at least one ear

Pure tone audiometry is measured separately in the left and right ear. The order of testing (left ear first, right ear first and alternating between ears) is recorded. Air conduction is tested, at 500, 1000, 2000 and 4000 Hz, and recorded as dB HL for each of the following:

- Left ear air conduction;
- Right ear air conduction.

Reasons for not conducting test are recorded as freetext and will be classified by the TOPS Lead Audiologist. If a child does not tolerate headphones for the Pure Tone Audiometry then the sound field audiometry should have been completed.

Sound field audiometry is tested at the same frequencies, at the same visit if Pure tone audiometry unsuccessful. Sound field audiometry will be used and interpreted in the same way as pure tone audiometry for participants who have not had pure tone audiometry assessed for either ear. Note sound field audiometry does not test both ears separately.

A minimum of one dB HL result at one frequency is required for inclusion in outcome comparison. The threshold to determine abnormal hearing levels depends on the tests completed.

For participants with pure tone audiometry tested; At any frequency, a threshold of 20 dB HL will determine if the patient has abnormal hearing levels (agreed in meeting with TOPS Lead Audiologist on 01/07/2019) such that:

- Values >20 dB HL will be categorised as “abnormal”;
- Values ≤20 dB HL will be categorised as “normal”.

A patient will be classified as “abnormal” for outcome analysis if at least one frequency indicates this.

For participants with Sound field audiometry: At any frequency, a threshold of 25 dB HL will determine if the patient has abnormal hearing levels (agreed in meeting with TOPS Lead Audiologist on 18/09/2019) such that:

- Values ≥25 dB HL will be categorised as “abnormal”;
- Values <25 dB HL will be categorised as “normal”.

A patient will be classified as “abnormal” for outcome analysis if at least one frequency indicates this.

A patient will be classified as “abnormal” pure tone audiometry in at least one ear (*Outcome ID:* SO_7bi_3y, SO_7bi_5y*, Appendix Table 23*) at three and five years of age if at least one frequency in at least one ear indicates this at the respective timepoint.

ii. Abnormal pure tone audiometry in both ears

This outcome will be derived using the same approach as one ear, see *Section 18.6.2.7.1*.

Participants must have had both ears tested separately to be included in the outcome set.

A patient will be classified as “abnormal” pure tone audiometry in both ears (*Outcome ID:* SO_7bii_3y, SO_7bii_5y*, Appendix Table 23*) at three and five years of age if both ears indicates this at the respective timepoint.

iii. Severity in better ear

The scores recorded for pure tone audiometry or, if missing, sound field audiometry, see *Section 18.6.2.7.1*, will be averaged. Participants must have had both ears tested to be included in the outcome set.

All averages will be based on four scores. When a score is missing at certain frequencies i.e. scores missing for one, two or three of the four frequencies, the missing value(s) must be imputed using the rules defined in Table 21.

Table 21: Calculating missing frequencies for Hearing Level outcomes

Number of frequencies provided	Frequency missing (Hz, indicated with X)				Approach to impute missing value(s)
	500	1000	2000	4000	
3	X				Impute with 1000Hz value. [1]
				X	Impute with 2000Hz value. [1]
		X			Impute with average of 500Hz and 2000Hz values. [1]
			X		Impute with average of 1000Hz and 4000Hz values. [1]
2	X	X			Impute with 2000Hz value. [2]
	X		X		Impute 500Hz with 1000Hz value. [1] Impute 2000Hz with average of 1000Hz and 4000Hz.[1]
	X			X	Impute 500Hz with 1000Hz value. [1] Impute 4000Hz with 2000Hz value.[1]
		X	X		Impute 1000Hz with average of 500Hz and 4000Hz. [2] Impute 2000Hz with average of 500Hz and 4000Hz. [2]
		X		X	Impute 1000Hz with average of 500Hz and 2000Hz. [1] Impute 4000Hz with 2000Hz value. [1]
			X	X	Impute 2000Hz with 1000Hz value. [2] Impute 4000Hz with 1000Hz value. [2]
1	X	X	X		Impute all values with non missing frequency. [1]
	X		X	X	Impute all values with non missing frequency. [1]
	X	X		X	Impute all values with non missing frequency. [1]
		X	X	X	Impute all values with non missing frequency. [1]

[1] As discussed in statistics meeting with TOPs Lead Audiologist on 01/07/2019

[2]As agreed with TOPs Lead Audiologist on 19/08/2019 (email)

Where scores are presented on both left and right ear, the better ear will be selected for analysis as the ear with the lowest average. This average will be rounded to the nearest whole number and score used to classify the severity of better ear according to the following pure tone hearing thresholds: [35]

- “Normal” if average is ≤20 dB HL
- “Mild“ if average is between 21 and 40 dB HL
- “Moderate“ if average is between 41 and 70 dB HL
- “Severe“ if average is between 71 and 95 dB HL
- “Profound“ if average is >95 dB HL

A patient will be classified as severity in the better ear (*Outcome ID: SO_7bii_3y, SO_7bii_5y, Appendix Table 23*) at three and five years of age respectively according to the classification at the respective timepoint.

18.6.2.7.2 Analysis

a) 12 months:

i. Abnormal transient optoacoustic emission (TEOAE)

The following will be reported overall and split by treatment arm. Summaries will be reported as frequencies and percentages, defined as follows:

- Number of participants randomised (i)
- Number of participants with audiology assessment (ii denominator i), by group
- Number of participants not tested (iii, denominator ii), with reasons not done summarised
- Number of participants included in secondary outcome set (iv, denominator ii), with reasons provided where not clear from categories above
- Number of participants classified as having:
    - Abnormal TEOAE (v, denominator iv)
    - Normal TEOAE (vi, denominator iv)

The relative risk[34] together with 95% confidence interval (95% CI) will be reported along with a two-sided p-value from a chi-square test.

In the circumstance that the “Abnormal TEOAE” or “Normal TEOAE “ categorisations for each group contains less than five participants, thereby raising concern over the appropriateness of a chi-square test, a Fisher’s exact test will be used.

No sensitivity analysis will be performed.

ii. Abnormal Sound field audiometry

The following will be reported overall and split by treatment. Summaries will be reported as frequencies and percentages, defined as follows:

- Number of participants randomised (i)
- Number of participants with audiology assessment (ii, denominator i), by group
- Number of participants not tested (iii, denominator ii), with reasons not done summarised
- Number of participants tested with sound field only (iv, denominator ii)

- Number of participants tested with VRA only (v, denominator ii)
- Number of participants tested with sound field and VRA (vi, denominator ii)
    - Number of participants with different results across tests (vii, denominator vi)
        - Line listings of results will be provided.
    - Number of participants with the same results across tests (viii, denominator vi)
- Number of participants included in secondary outcome set (ix, denominator ii), with reasons provided where not clear from categories above
- Number of frequencies with results per patient (x, denominator ix)
- Number of participants classified as having:
    - Abnormal sound field audiometry (xi, denominator ix)
    - Normal sound field audiometry (xii, denominator ix)

The mean sound field audiometry result will be computed following the rules defined in *Section 18.6.2.7.1* and Table 21, and summarised by means and standard deviations (SDs), medians and inter-quartile range (IQR) if data are skewed. Minimum and maximum values will also be presented. Summaries will be presented both overall, and split by treatment arm.

For each time point, the relative risk[34] together with 95% confidence interval (95% CI) will be reported along with a two-sided p-value from a chi-square test.

In the circumstance that the "Abnormal sound field audiometry" or "Normal sound field audiometry" categorisations for each group contains less than five participants, thereby raising concern over the appropriateness of a chi-square test, a Fisher's exact test will be used.

No sensitivity analysis will be performed.

b) 3 and 5 years

i. Abnormal pure tone audiometry in at least one ear

The following will be reported for each timepoint both overall and split by treatment arm. Summaries will be reported as frequencies and percentages, defined as follows:

- Number of participants randomised (i)
- Number of participants with audiology assessment (ii, denominator i), by group
- Number of participants not tested (iii, denominator ii), with reasons not done summarised
- Number of participants tested with pure tone only (iii, denominator ii)
    - One ear tested (iv, denominator iii)

    - Both ears tested (v, denominator iii)
- Number of participants tested with sound field only (vi, denominator ii)
- Number of participants tested with pure tone and sound field (vii, denominator ii)
    - One ear tested (viii, denominator vii)
    - Both ears tested (ix, denominator vii)
    - Number of participants with different results across tests (x, denominator vii)
        - Line listings of results will be provided.
    - Number of participants with the same results across tests (xi, denominator vii)
- Number of participants included in secondary outcome set (xii, denominator ii), with reasons provided where not clear from categories above
- Number of frequencies with results per patient (xiii, denominator xii)
- Number of participants classified as having:
    - Abnormal pure tone audiometry in at least one ear (xiv, denominator xii)
    - Abnormal pure tone audiometry in neither ear (xv, denominator xii)

The mean audiometry result will be computed following the rules defined in *Section 18.6.2.7.1* and Table 21, and summarised by means and standard deviations (SDs). Minimum and maximum values will also be presented. The mean will then be used to classify each child as abnormal/normal hearing levels in at least one ear and frequencies and percentages provided (denominator xii). Summaries will be presented by timepoint both overall, and split by treatment (denominator xii overall, by timepoint and each treatment group).

The relative risk[34] together with 95% confidence interval (95% CI), according to the classification in *Section 18.6.2.7.1,* will be reported along with a two-sided p-value from a chi-square test.

In the circumstance that the “Abnormal pure tone audiometry in at least one ear” or “Abnormal pure tone audiometry in neither ear“ categorisations for each group contains less than five participants, thereby raising concern over the appropriateness of a chi-square test, a Fisher’s exact test will be used.

No sensitivity analysis will be performed.

ii. Abnormal pure tone audiometry in both ears

This outcome will be analysed following the same approach as one ear, see *Section 18.6.2.7.2*, with abnormal pure tone audiometry in both ears being the outcome of interest reported and analysed.

iii. Severity of better ear (normal, mild, moderate, severe, profound)

This outcome will be summarised as part of the analysis of abnormal pure tone audiometry in one or both ears, see *Section 18.6.2.7.2*.

In addition, the following will be reported for each timepoint both overall and split by treatment arm. Summaries will be reported as frequencies and percentages (denominator xii), defined as follows:

- Number of participants classified as having severity of better ear as:
    - “Normal” (≤20 dB HL)
    - “Mild“ (between 21 and 40 dB HL)
    - “Moderate“ (between 41 and 70 dB HL)
    - “Severe“ (between 71 and 95 dB HL)
    - “Profound“ (>95 dB HL)

For each time point, a two-sided p-value from a chi-square test for trend will be reported. In the circumstance that the numbers within severity categorisations for each group contains less than five participants, thereby raising concern over the appropriateness of a chi-square test for trend, groups will be combined to ensure suitability e.g. participants with severe or profound severity combined as one group. If this is not possible a proportional odds model will be applied.

No sensitivity analysis will be performed.

18.6.2.8 Secondary outcome 8: Middle ear function

This outcome is measured using the following end points at 12 months, 3 years and 5 years:

a) Flat line tympanogram in at least one ear
b) Flat line tympanogram in both ears

18.6.2.8.1 Derivation

Middle ear function at 12 months is collected on *Form 6: 12 months audiology data collection.* Middle ear function at 3 and 5 years is collected on *Form 8: Audiology data collection 3 and 5 years.* Both forms record the same data for deriving the middle ear function endpoints at set timepoints.

a) Flat line tympanogram in at least one ear

Tympanic membrane is classified as Type A, Type B, or Type C. Both ears are tested separately and results recorded. Reasons for "Not tested" should be provided as freetext and will be classified by the TOPS Lead Audiologist. If a response is missing it will be assumed the child is not tested, inconclusive or cannot be interpreted.

Flat line tympanogram in at least one ear *(Outcome ID: SO_8a_12m, SO_8a_3y, SO_8a_5y, Appendix Table 23)* is defined as a response of "Type B" to the tympanic membrane test for either ear tested.

The ear canal volume of each ear tested is also collected and can provide insight into the validity of a type B (flat) response. A threshold equivalent ear canal volume (ECV) 0.30 ml will be used to determine a blockage, with a blockage being classified as <0.30 ml (agreed in meeting with TOPS Lead Audiologist on 18/09/2019). All participants with a blockage will be reported according to the classification of type of tympanic membrane they present.

Update following blind review:
The classification according to the worst ear would use Type B, Type C then Type A, whereby B is the worst classification and A the better.

b) Flat line tympanogram in both ears

This outcome will be derived using the same approach as one ear, see *Section 18.6.2.8.1*.

Participants must have had both ears tested separately to be included in outcome set.

A patient will be classified as having flat line tympanogram in both ears (*Outcome ID: SO_8b_12m, SO_8b_3y, SO_8b_5y, Appendix Table 23*) at 12 months, three years and five years if both ears indicates this at the respective timepoint.

18.6.2.8.2 Analysis

a) Flat line tympanogram in at least one ear

The following will be reported for each timepoint both overall and split by treatment arm. Summaries will be reported as frequencies and percentages, defined as follows:

- Number of participants randomised (i)
- Number of participants with audiology assessment (ii, denominator i), by group
- Number of participants not tested for tympanic membrane (iii, denominator ii), with reasons not done summarised
- Number of participants included in secondary outcome set (iv denominator ii), with reasons provided where not clear from categories above
- Number of participants tested for tympanic membrane (v, denominator iv)
    - One ear tested (vi, denominator v)
    - Both ears tested (vii, denominator v)
- Number of participants classified as tympanic membrane according to the worst ear:
    - Type A (viii, denominator v)
    - Type C (ix, denominator v)
    - Type B (x, denominator v)
    - Functioning Ventilation tube/Perforation only indicated (xi, denominator v)
- Number of participants with a tympanic membrane, with an indication of Functioning Ventilation tube/Perforation on the CRF (xii) and:
    - Type A and/or Type C (xiii, denominator xii)
    - Type B (xiv, denominator xii)
    - No type indicated (xv, denominator xii)
- Number of participants with blockage indicated in outcome ear (xvi, denominator v), by type
    - Less than 0.30 ml (xvii, denominators v,viii,ix,x,xi)
    - Greater than or equal to 0.30 ml (xviii, denominators v,viii,ix,x,xi)
- Number of participants classified as having a tympanogram Type A, B or C (xixi) classified as:
    - Flat line tympanogram in at least one ear (xx, denominator xix)
    - Flat line tympanogram in neither ear (xxi, denominator xix)

For each time point the relative risk[34] together with 95% confidence interval (95% CI), according to the classification in *Section 18.6.2.8.1,* will be reported along with a two-sided p-value from a chi-square test.

In the circumstance that the "Flat line tympanogram in at least one ear" or "Flat line tympanogram in neither ear" categorisations for each group contains less than five participants, thereby raising concern over the appropriateness of a chi-square test, a Fisher's exact test will be used.

No sensitivity analysis will be performed.

b) Flat line tympanogram in both ears

This outcome will be analysed following the same approach as one ear, see *Section 18.6.2.8.2*, with flat line tympanogram in both ears being the outcome of interest reported and analysed.

18.6.2.9 Secondary outcome 9: Dentofacial development at five years

This outcome is measured using the following end points:

a) Soft tissue ANB (the angle between soft tissue nasion: A point, and B point on profile photograph)
b) Maxillary arch constriction score (using modified Huddart/Bodenham scoring system)

18.6.2.9.1 Derivation

a) Soft tissue ANB (the angle between soft tissue nasion: A point, and B point on profile photograph)

Soft tissue ANB (*Outcome ID: SO_9a_5y, Appendix Table 23*) is measured on child's left or right, where left not available, profile photographs taken at five years of age. The photographs are scanned and the angle measured using the validated Dolphin Imaging Plus software. The software output is databased using *Form 16: Soft Tissue ANB*. The angle is expected to be less than 30 degrees and is measured to one decimal place.

Two assessors will be needed to assess the photographs as agreed in the TMG 15/08/2019. All photographs are to be measured by one assessor and approximately 20% repeated for intra-rater assessment. A second assessor will assess at least 50 of the photographs in the first instance for inter-rater assessment, if results of inter rater analysis indicates the need to the number of photographs requiring a second reviewer will be increased. Angles measured by the assessor who rates the whole cohort will be used for the outcome assessment. For angles measured twice by the same assessor, the second angle will be used for outcome assessment.

The endpoint soft tissue ANB is the angle as recorded on *Form 16: Soft Tissue ANB.*

Free text reasons for failing photograph quality checks will be categorised by the Dentofacial lead.

b) Maxillary arch constriction score (using modified Huddart/Bodenham scoring system)

Maxillary arch constriction score (*Outcome ID: SO_9b_5y, Appendix Table 23*) is measured on maxillary arch impressions taken at five years of age. The impressions are scanned to enable their digital storage for endpoint assessment. The score is measured using the Landmark Recognition software. No manual scoring of the impressions are necessary (agreed in the TMG 15/08/2019). The software output is databased using *Form 17: Maxillary arch constriction score.* The score ranges from -24 to 8 and is measured to whole numbers.

All impressions are to be scored by one assessor and 20% repeated for intra-rater assessment. A second assessor will assess 50 of the TOPS impressions for inter-rater assessment (agreed in the meeting with AC and Dundee 19/08/2019 and TMG 15/08/2019). Impression scores from the same assessor will be used for the outcome assessment. A proportion of models are to use a manual score rather than the software score as the software had problems identifying the landmarks in a number of models.

The endpoint maxillary arch contriction score is the score as recorded on *Form 17: Maxillary arch contriction score.*

Free text reasons for failing impression quality checks will be categorised by the Dentofacial lead.

18.6.2.9.2 Analysis

a) Soft tissue ANB (the angle between soft tissue nasion: A point, and B point on profile photograph)

The following will be reported overall and split by treatment arm. Summaries will be reported as frequencies and percentages, defined as follows:

- Number of participants randomised (i)
- Number of participants with photograph taken, reasons for missing data will be provided
    - Photograph (ii, denominator i)
    - No photograph (iii, denominator i)
- Number of participants with photograph taken by site.
- Number of participants with photographs eligible for assessment:
    - Adequate quality for assessment (iv, denominator ii)
    - Insufficient quality for assessment (v, denominator ii)
- Number of participants included in secondary outcome set (vi, denominator iv), with reasons provided where not clear from categories above
- Number of participants with the side of photograph used:
    - Right (vii, denominator vi)
    - Left (viii, denominator vi)

Soft tissue ANB will be summarised overall, and split by treatment group, by means and standard deviations (SDs), medians and inter-quartile range (IQR) if data are skewed. Minimum and maximum values will also be presented.

If the data is normally distributed (assessed using QQ plot by treatment group), the mean soft tissue ANB will be tested using a t-test, and the mean difference and 95% CI reported. Otherwise, the mean score will be tested using a non-parametric equivalent test.

No sensitivity analysis will be performed.

Update following blind review:
Data is normally distributed for the outcome therefore, soft tissue ANB score will be summarised by means and SDs. The mean score will be tested using a t-test.

b) Maxillary arch constriction score (using modified Huddart/Bodenham scoring system)
This outcome will be analysed following the same approach soft issues ANB, see *Section 18.6.2.9.2*, with maxillary arch constriction score being the outcome of interest reported and analysed.

In addition, the number of participants with each of the individual 33 scores (-24 to 8) will be reported as percentages (denominator vi) and frequencies.

Update following blind review:
Data is normally distributed for the outcome therefore, maxillary arch constriction score will be summarised by means and SDs. The mean score will be tested using a t-test.

18.6.2.9.2.1 Inter-rater and Intra-rater analysis

Inter-rater and intra-rater analysis will be performed using data from photographs rated more than once using Bland Altman agreement analysis.[36]

The limits of agreement of the soft tissue ANB is 1.5. The limit of agreement of maxilliary arch constriction score is 3.

18.6.2.10 Secondary outcome 10: Growth at 12 months

This outcome is measured using the following end points:

a) Nude weight
b) Crown to heel length
c) Occipitofrontal circumference

18.6.2.10.1 Derivation

a) Nude weight

Nude weight at 12 months is collected on *Form 5: 12 month growth assessment*.

Nude weight *(Outcome ID: SO_10a_12m, Appendix Table 23)* is measured in grams to the nearest whole number.

The endpoint nude weight is the weight as recorded on *Form 5: 12 month growth assessment.* Reasons for "Not recorded" should be provided as freetext and will be classified by the CInv, or delegated other.

b) Crown to heel length

Crown to heel length at 12 months is collected on *Form 5: 12 month growth assessment*.

Crown to heel length *(Outcome ID: SO_10b_12m, Appendix Table 23)* is measured in centimeters (cm) to one decimal place.

The endpoint crown to heel length is the length as recorded on *Form 5: 12 month growth assessment.* Reasons for "Not recorded" should be provided as freetext and will be classified by the CInv, or delegated other.

c) Occipitofrontal circumference

Occipitofrontal circumference at 12 months is collected on *Form 5: 12 month growth assessment*.

Occipitofrontal circumference *(Outcome ID: SO_10c_12m, Appendix Table 23)* is measured in centimeters (cm) to one decimal place.

The endpoint occipitofrontal circumference is the length as recorded on *Form 5: 12 month growth assessment.* Reasons for "Not recorded" should be provided as freetext and will be classified by the CInv, or delegated other.

18.6.2.10.2 Analysis

a) Nude weight

The following will be reported overall and split by treatment arm. Summaries will be reported as frequencies and percentages, defined as follows:

- Number of participants randomised (i)
- Number of participants with growth assessment (*Form 5: 12 month growth assessment)*, reasons for missing data will be provided
    - Growth assessment (ii, denominator i)
    - No growth assessment (iii, denominator i)
- Number of participants with nude weight eligible for assessment:
    - Nude weight measured (iv, denominator ii)
    - Nude weight not measured (v, denominator ii)
- Number of participants included in secondary outcome set (vi denominator iv), with reasons provided where not clear from categories above

Nude weight will be summarised overall, and split by treatment group, by means and standard deviations (SDs), medians and inter-quartile range (IQR) if data are skewed. Minimum and maximum values will also be presented.

If the data is normally distributed (assessed using QQ plot by treatment group), the mean nude weight will be tested using a t-test, and the mean difference and 95% CI reported. Otherwise, the mean score will be tested using a non-parametric equivalent test.

No sensitivity analysis will be performed.

Update following blind review:
Data is normally distributed for the outcome therefore, nude weight will be summarised by means and SDs. The mean score will be tested using a t-test.

b) Crown to heel length

This outcome will be analysed following the same approach as nude weight, see *Section 18.6.2.10.2*, with crown to heel length being the outcome of interest reported and analysed.

Update following blind review:

Data is normally distributed for the outcome therefore, crown to heel length will be summarised by means and SDs. The mean score will be tested using a t-test.

c) Occipitofrontal circumference

This outcome will be analysed following the same approach as nude weight, see *Section 18.6.2.10.2*, with occipitofrontal circumstance being the outcome of interest reported and analysed.

Update following blind review:

Data is normally distributed for the outcome therefore, occipitofrontal circumference will be summarised by means and SDs. The mean score will be tested using a t-test.

19 Missing data and withdrawals

Based on experience from the ScandCleft study the structure of the centralized cleft palate care system and trial specific systems in place will ensure the occurrence of missing data is likely to be low. Therefore the potential impact of any missing data is likely to be low.

In the primary analysis the principle of intention to treat will be used as far as practically possible. If required, see *Section 18.1*, a per protocol analysis will also be conducted. Unless specified otherwise in the analysis sections, neither approach will contain imputed data.

Participants with insufficient data to make their assessments will be expressed as a frequency and a percentage with the denominator being those who were randomised, treated and consented.

Withdrawals over the course of the trial will be summarised by treatment arm. This will be presented as described in *Section 14.1* and *14.2.*

20 Additional analyses

20.1 DENVER II

The DENVER II test is collected at 3 years on *Form 10: DENVER II assessment at three years*. The following will be reported overall and split by treatment. Summaries will be reported as frequencies and percentages, defined as follows:

- Number of participants randomised (i)
- Number of participants with Form 10 complete (ii, denominator i)
- Interpretation of DENVER II
    - Normal [no delays and a maximum of one caution] (iii, denominator ii)
    - Suspect [two or more cautions and/or one or more delays] (iv, denominator ii)
    - Untestable [refusal scores on one or more items] (v, denominator ii)
- Severe developmental delay diagnosed
    - Yes (vi, denominator ii)
    - No (vii, denominator ii)

20.2 Speech therapy outside of routine trial visits

Treatment for speech therapy outside of routine visits is collected at 12 months, 3 years and 5 years.

Data at 12 months, collected on *Form 4: 12 month speech assessment* (12 months), asks whether or not a child is seen by a speech and language therapist since birth and whether or not the therapist is based at trial site. The number of sessions is also provided.

Data at 3 years, collected on *Form 9: 3 Year Speech Assessment* and *Form 11: 3 year speech therapy questionnaire*, asks whether or not a child is seen by a speech and language therapist since 12 month assessment and whether or not the therapist is based at trial site. The number of sessions is also provided.

Data at 5 years, collected on *Form 13: 5 Year Speech Assessment* and *Form 14: 5 year speech therapy questionnaire*, asks whether or not a child is seen by a speech and language therapist since 3 year assessment and whether or not the therapist is based at trial site. The number of sessions is also provided.

The following will be reported overall and split by treatment arm and trial site. Summaries will be reported as frequencies and percentages, defined as follows:

- Number of participants randomised (i)

- Number of participants with Form 4 complete (ii, denominator i)
    - Seen by speech and language therapist based at trial site since birth [1] (iii, denominator ii)
        - Received counselling/advice to parents [1] (iv, denominator iii)
        - Received counselling/advice to others [1] (v, denominator iii)
        - Received treatment speech therapy – information [1] (vi, denominator iii)
        - Received treatment speech therapy – early intervention [1] (vii, denominator iii)
- Number of participants with Form 9 or Form 11 complete (viii, denominator i)
    - Seen by any speech and language therapist (local or trial site based) [2,3] (ix, denominator viii)
    - Seen by speech and language therapist based at trial site since 12 month [2] (x, denominator viii)
        - Reviewed/monitored [1,2] (xi, denominator x)
        - Received direct therapy [1,2] (xii, denominator x)
        - Received indirect therapy [1,2] (xii, denominator x)
    - Seen by a local speech and language therapist since birth[2,3] (xiii, denominator viii)
        - Reviewed/monitored [1,3] (xiv, denominator xiii)
        - Received direct therapy [1,3] (xv, denominator xiii)
        - Received indirect therapy [1,3] (xvi, denominator xiii)
- Number of participants with Form 13 or Form 14 complete (xvii, denominator i)
    - Seen by any speech and language therapist (local or trial site based) [4,5] (xviii, denominator xvii)
    - Seen by speech and language therapist based at trial site since 3 year[1,4] (xix, denominator xvii)
        - Reviewed/monitored [4] (xx, denominator xix)
        - Received direct therapy [1,4] (xxi, denominator xix)
        - Received indirect therapy [1,4] (xxii, denominator xix)
    - Seen by a local speech and language therapist since last seen for TOPs [4,5] (xxiii, denominator xvii)
        - Reviewed/monitored [1,5] (xxiv, denominator xxiii)
        - Received direct therapy [1,5] (xxv, denominator xxiii)
        - Received indirect therapy [1,5] (xxvi, denominator xxiii)

Notes:

[1] Present number of times by means and standard deviations (SDs), medians and inter-quartile range (IQR) if data are skewed. Minimum and maximum values will also be presented.
[2] Source: Form 9
[3] Source: Form 11
[4] Source: Form 13
[5] Source: Form 14

20.3 Nasometry

Nasometry is collected at 5 years on *Form 13: Five year speech assessment*. The following will be reported overall and split by treatment arm and primary language of recruiting site (Brazil, Denmark, Norway, Sweden, UK). Summaries will be reported as frequencies and percentages, defined as follows:

- Number of participants randomised (i)
- Number of participants with nasometry assessment (ii, denominator i)
- Number of sequences provided per patient (1, 2, 3, 4, 5) (denominator ii)

Nasometry across the five sequences will be summarised overall and split by treatment arm and primary language, by means and standard deviations (SDs), medians and inter-quartile range (IQR) if data are skewed. Minimum and maximum values will also be presented.

If the data is normally distributed (assessed using QQ plot by treatment group), the mean nasometry will be tested using a t-test, and the mean difference and 95% CI reported. Otherwise, the mean score will be tested using a non-parametric equivalent test.

As nasometry is an acoustic correlate to hypernasality (*Section 18.5.2.3*), a box plot will be presented to show the nasometry scores of patients against the hypernasity scores of patients. All patients with both readings will be included in this set.

The approach to summarising the five sequences was agreed with the CSG on 12/12/2018.

Update following blind review:
Data is skewed for the outcome therefore, nasometry will be summarised by medians and IQR and analysed using a Mann Whitney U test.

20.4 Secondary surgery

Secondary surgeries are collected, as required, on *Form 12: Secondary surgeries*. The following will be reported overall and split by treatment. Summaries will be reported as frequencies and percentages, defined as follows:

- At patient level
    - Number of participants randomised (i)
    - Number of participants with Form 12 complete (ii, denominator i)
    - Number of secondary surgerys per patient (1, 2, 3, 4+) (iii, denominator ii)
    - Age at first secondary surgery (iv)
- At surgery level
    - Age at surgery
    - Reason for secondary surgery
        - Dehiscence (v, denominator iii)
        - Fistula (vi, denominator iii)
        - Velopharangeal insufficiency (vii, denominator iii)
        - Other (viii, denominator iii)
    - What was done at surgery
        - Fistula closure (ix, denominator iii)
        - Re-repair (x, denominator iii)
        - Pharyngoplasty (xi, denominator iii)
        - Other (xii, denominator iii)

Where applicable, these details will be summarised as line listings for participants with multiple secondary surgeries.

21 Safety evaluations

21.1 Data sets analysed

Reporting non-serious safety observations

When reporting baseline surgery characteristics, intra operative events, early complications during the hospital stay, observations monitored 48 hours post surgery, post operative medication and late complications, see *Section 21.2.1* to *Section 21.2.6,* the principle of intention-to-treat (ITT) will be the main strategy of the analysis. These analyses will be conducted on all randomised participants, in the group to which they were allocated, and for whom the observations of interest have been observed/measured.

Reporting Serious Adverse Events and Unanticipated Problems

When reporting Serious Adverse Events (SAEs) or Unanticipated Problems, see *Section 21.2.7* and *Section 21.2.8* respectively, participants will be analysed according to their date of surgery. This is to accurately represent the events with respect to the intervention itself i.e. surgical timing. All participants who receive surgery will be included in the safety analysis population and classified according to the following groups:

- 6 month safety group: receives surgery prior to 9 months of age (corrected for gestational age)
- 12 month safety group: receives surgery on or after 9 months of age (corrected for gestational age)

The numbers within the safety data set will be presented as count data and as a percentage using the total randomised as a denominator. This will be presented both overall and split by treatment group.

A scatterplot and/or bubble chart, with an x-axis of actual surgery timing and y-axis of randomised surgery group, will be presented.

21.2 Presentation of the data

Safety events will be presented using descriptive statistics.

Tests of statistical significance will not be undertaken for safety events; rather the clinical importance of any imbalance will be noted.

Details to the presentation of each event type are provided in the following sections.

21.2.1 Baseline surgery characteristics

Baseline surgery characteristics are collected on *Form 2: Surgery and Postoperative Care*. The following will be reported overall and split by treatment arm, defined as follows:

- Grading of cleft at surgery * (Grade 1, 2, 3 or 4)
- Shape of cleft *
    - “U” shaped
    - "V” shaped
- Dimension of cleft palate;
    - Soft tissue width at posterior hard palate ^ (mm)
    - Bony width at posterior hard palate ^ (mm) [1]
    - Width at base of uvula ^ (mm)
    - Length of soft palate (distal base of uvula-hard palate) ^ (mm)

Categorical variables (indicated *) will be reported as frequencies and percentages. Continuous variables (indicated ^) will be reported by means and standard deviations (SDs), medians and inter-quartile range (IQR) if data are skewed. Minimum and maximum values will also be presented. [1]Some surgeons have taken the measurement incorrectly as the hard palate bony shelves, these will be removed from ‘bony width at posterior hard palate’ summary and summarised seperatley.

In addition, the randomisation process is minimised on the stratification factor size of cleft:

- Soft palate at randomisation should be classified as grades 1 or 2 at surgery
- Soft and hard palate should be classified as grades 3 and 4 at surgery.

The number of participants whose categorisation changes between randomisation and surgery will be summarised both overall, and split by treatment group, by frequencies and percentages.

21.2.2 Intra operative events

Intra operative events are collected on *Form 2: Surgery and Postoperative Care*. The following will be reported overall and split by treatment. Summaries will be reported as frequencies and percentages, defined as follows:

- Blood transfusion during surgery (yes/no)
- Anaesthetic complications (yes/no)
- Bleeding (yes/no)

21.2.3 Early complications during hospital stay

Early complications during hospital stay are collected on *Form 2: Surgery and Postoperative Care*. The following will be reported overall and split by treatment. Summaries will be reported as frequencies and percentages, defined as follows:

- Postoperative airway problems (yes/no)
- Postoperative blood loss (prior to discharge from hospital) (yes/no)
- Anticoagulants given (yes/no)
- Dehiscence (yes/no)
- Readmission to operating room (yes/no)

21.2.4 Observations monitored 48 hours post surgery

Observations monitored 48 hours post surgery are collected on *Form 2: Surgery and Postoperative Care*. The following will be reported overall and split by treatment. Summaries will be reported as frequencies and percentages, defined as follows:

- Oxygen saturation levels
    - Clinically significant abnormality
    - No clinically significant abnormality
    - No observation
- Carbon dioxide and oxygen
    - Clinically significant abnormality
    - No clinically significant abnormality
    - No observation
- Arterial blood gases
    - Clinically significant abnormality
    - No clinically significant abnormality
    - No observation
- Heart rate
    - Clinically significant abnormality/treatment required
    - Clinically significant abnormality/no treatment required
    - No clinically significant abnormality
    - No observation
- Blood pressure
    - Clinically significant abnormality/treatment required
    - Clinically significant abnormality/no treatment required
    - No clinically significant abnormality
    - No observation

- Respiration
    - Clinically significant abnormality/treatment required
    - Clinically significant abnormality/no treatment required
    - No clinically significant abnormality
    - No observation
- Body temperature
    - Clinically significant abnormality/treatment required
    - Clinically significant abnormality/no treatment required
    - No clinically significant abnormality
    - No observation

21.2.5 Post operative medication

Post operative medication are collected on *Form 2: Surgery and Postoperative Care*. The following will be reported overall and split by treatment. Summaries will be reported as frequencies and percentages, defined as follows:

- Pain relief medication (yes/no)
- Anti-inflammatory medication (yes/no)
- Antibiotic medication (other than prophylactic antibiotics) (yes/no)

21.2.6 Late complications

Late complications are collected on *Form 3: Late complications*. The following will be reported overall and split by treatment. Summaries will be reported as frequencies and percentages, defined as follows:

- Secondary bleeding (yes/no)

21.2.7 Serious adverse events

Serious adverse events (SAE) are collected on *Form A1: Serious Adverse Events*. Serious adverse events will be presented as line listings including the following information:

- SAE number
- Centre
- Treatment group
- Main diagnostic symptom (freetext classified by CInv, or delegated other)
- Severity (mild/moderate/severe)
- Seriousness
    - Principal Investigator (PI) classification
        - Resulted in death

        - Life-threatening
        - required inpatient hospitalisation
        - Prolonged existing hospitalisation
        - Persistent or significant disability/Incapability
        - Other important medical condition
    - CInv classification
        - Resulted in death
        - Life-threatening
        - required inpatient hospitalisation#
        - Prolonged existing hospitalisation
        - Persistent or significant disability/Incapability
        - Other important medical condition
- Relationship to intervention
    - Principal Investigator (PI) classification
        - Unrelated
        - Unlikely
        - Possibly
        - Probably
        - Almost certainly
    - CInv classification
        - Unrelated
        - Unlikely
        - Possible
        - Probably
        - Definitely
- Expectedness
    - Principal Investigator (PI) classification (expected/unexpected)
    - CInv classification (expected/unexpected)
- If unrelated or unlikely, most likely cause classified by PI
    - Other illness
    - Prior or concomitant treatment
    - Protocol procedure
    - Other
- Unanticipated problem
    - Principal Investigator (PI) classification (yes/no)
    - CInv classification (yes/no)

- Action taken
    - Readmission to the operating theatre
    - Prolongation of hospital stay
    - Treatment with medication
    - Hospital readmission
    - Attendance at hospital emergency department
    - Attendance to general practitioner or other medical care team
    - Other
- Outcome
    - Resolved
    - Resolved with sequelae
    - Not resolved/Ongoing
    - Ongoing at final follow up
    - Fatal
    - Unknown
- Patient status
    - Continue follow up
    - Withdrawn from follow up

21.2.8 Unanticipated problems

Unanticipated problems are collected on *Form U1: Unanticipated problems*. Unanticipated problems will be presented as line listings including the following information:

- SAE number (if also an adverse event)
- Centre
- Treatment group
- Detailed description of event (freetext)
- Description of why an unanticipated problem (freetext)
- Action taken (freetext)

22 References

1. Tanaka SA, Mahabir RC, Jupiter DC, et al. Updating the Epidemiology of Isolated Cleft Palate. Plastic and reconstructive surgery 2013;131(4):650e-52e doi: 10.1097/PRS.0b013e3182827790[published Online First: Epub Date]|.
2. Peterson-Falzone S. Optimal age for palatoplasty to faclitate normal speech development: What is the evidence? In: Berkowitz S, ed. Cleft Lip and Palate. Berlin: Springer, 2006:691-700.
3. Hellquist R, Ponten B, Skoog T. The influence of cleft length and palatoplasty on the dental arch and the deciduous occlusion in cases of clefts of the secondary palate. Scandinavian journal of plastic and reconstructive surgery 1978;12(1):45-54
4. Friede H, Persson EC, Lilja J, et al. Maxillary dental arch and occlusion in patients with repaired clefts of the secondary palate. Influence of push back palatal surgery. Scandinavian journal of plastic and reconstructive surgery and hand surgery / Nordisk plastikkirurgisk forening [and] Nordisk klubb for handkirurgi 1993;27(4):297-305
5. Nystrom M, Ranta R. Effect of timing and method of closure of isolated cleft palate on development of dental arches from 3 to 6 years of age. European journal of orthodontics 1994;16(5):377-83
6. Friede H, Enocson L, Moller M, et al. Maxillary dental arch and occlusion in repaired clefts of the secondary palate: influence of surgical closure with minimal denudation of bone. Scandinavian journal of plastic and reconstructive surgery and hand surgery / Nordisk plastikkirurgisk forening [and] Nordisk klubb for handkirurgi 2000;34(3):213-8
7. Dorf DS, Curtin JW. Early cleft palate repair and speech outcome. Plastic and reconstructive surgery 1982;70(1):74-81
8. Chapman KL, Hardin-Jones MA, Goldstein JA, et al. Timing of palatal surgery and speech outcome. The Cleft palate-craniofacial journal : official publication of the American Cleft Palate-Craniofacial Association 2008;45(3):297-308 doi: 10.1597/06-244[published Online First: Epub Date]|.
9. Chapman KL, Willadsen E. The Development of Speech in Children with Cleft Palate. Cleft Palate Speech: Assessment and Intervention 2011:23-40 doi: Book_Doi 10.1002/9781118785065[published Online First: Epub Date]|.
10. Jones DL. Timing of palatoplasty and speech. In: Losee JE, ed. Comprehensive Cleft Care: Taylor & Francis Group, LLC, 2016:521-26.[1
11. Semb G, Enemark H, Friede H, et al. Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 1. Planning and management. Journal of plastic surgery and hand surgery 2017;51(1):2-13 doi: 10.1080/2000656X.2016.1263202[published Online First: Epub Date]|.
12. Lohmander A, Willadsen E, Persson C, et al. Methodology for speech assessment in the scandcleft project - An international randomized clinical trial on palatal surgery: Experiences from a pilot study. Cleft Palate-craniofacial Journal 2009;46(4):347-62
13. Bannister P, Lindberg N, Jeppesen K, et al. Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 3. Descriptive study of postoperative nursing care following first stage cleft closure. Journal of plastic surgery and hand surgery 2017;51(1):21-26 doi: 10.1080/2000656X.2016.1269776[published Online First: Epub Date]|.
14. Feragen KB, Rumsey N, Heliovaara A, et al. Scandcleft randomised trials of primary surgery for unilateral cleft lip and Palate: 9. Parental report of social and emotional experiences related to their 5-year-old child's cleft diagnosis. Journal of plastic surgery and hand surgery 2017;51(1):73-80 doi: 10.1080/2000656X.2016.1254643[published Online First: Epub Date]|.
15. Feragen KB, Semb G, Heliovaara A, et al. Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 10. Parental perceptions of appearance and treatment outcomes in their 5-year-old child. Journal of plastic surgery and hand

surgery 2017;51(1):81-87 doi: 10.1080/2000656X.2016.1254642[published Online First: Epub Date]|.

16. Heliovaara A, Kuseler A, Skaare P, et al. Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 6. Dental arch relationships in 5 year-olds. Journal of plastic surgery and hand surgery 2017;51(1):52-57 doi: 10.1080/2000656X.2016.1221352[published Online First: Epub Date]|.

17. Karsten A, Marcusson A, Hurmerinta K, et al. Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 7. Occlusion in 5 year-olds according to the Huddart and Bodenham index. Journal of plastic surgery and hand surgery 2017;51(1):58-63 doi: 10.1080/2000656X.2016.1265529[published Online First: Epub Date]|.

18. Lohmander A, Persson C, Willadsen E, et al. Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 4. Speech outcomes in 5-year-olds - velopharyngeal competency and hypernasality. Journal of plastic surgery and hand surgery 2017;51(1):27-37 doi: 10.1080/2000656X.2016.1254645[published Online First: Epub Date]|.

19. Molsted K, Humerinta K, Kuseler A, et al. Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 8. Assessing naso-labial appearance in 5-year-olds - a preliminary study. Journal of plastic surgery and hand surgery 2017;51(1):64-72 doi: 10.1080/2000656X.2016.1266492[published Online First: Epub Date]|.

20. Rautio J, Andersen M, Bolund S, et al. Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 2. Surgical results. Journal of plastic surgery and hand surgery 2017;51(1):14-20 doi: 10.1080/2000656X.2016.1254646[published Online First: Epub Date]|.

21. Shaw W, Semb G. The Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 11. What next? Journal of plastic surgery and hand surgery 2017;51(1):88-93 doi: 10.1080/2000656X.2016.1254644[published Online First: Epub Date]|.

22. Willadsen E, Lohmander A, Persson C, et al. Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 5. Speech outcomes in 5-year-olds - consonant proficiency and errors. Journal of plastic surgery and hand surgery 2017;51(1):38-51 doi: 10.1080/2000656X.2016.1254647[published Online First: Epub Date]|.

23. Bearn D, Sandy JR and Shaw, WC. Photogrammetric assessment of the soft tissue profile in unilateral cleft lip and palate. Cleft palate-Craniofacial J. 2002; 39, 597-603.

24. Gray D and Mossey PA. Evaluation of a modified Huddart/Bodenham system for assessment of maxillary constriction. European Journal of Orthodontics 2005; 27, 507-11.

25. Martin CB, Ma X, McIntyre GT, et al. The validity and reliability of an automated method of scoring dental arch relationships in unilateral cleft lip and palate using the modified Huddart-Bodenham scoring system. Eur J Orthod 2016;38:353–8.

26. Sommerlad BC. A technique for cleft palate repair. Plastic and reconstructive surgery 2003;112(6):1542-8 doi: 10.1097/01.PRS.0000085599.84458.D2[published Online First: Epub Date]|.

27. Sommerlad BC. A technique for cleft palate repair. Plast Reconstr Surg. 2003; 112: 1542-1548.

28. Sommerlad BC. Comprehensive cleft care. 1st Edition. McGraw Hill Medicine. 2009.

29. Jensen BL, Kreiborg S, Dahl E, Fogh-Andersen P. Cleft Lip and Palate in Denmark, 1976-1981: Epidemiology, Variability, and Early Somatic Development. Cleft Palate Journal 1988; (25) 3: 258-269

30. NIH, Toolkit - Interventional Studies, March 2019 https://www.nidcr.nih.gov/research/human-subjects-research/interventional-studies/planning-and-start-up?_ga=2.96526620.288960275.1559559091-1372681346.1516700384

31. International Conference on Harmonisation. Topic E6 Good Clinical Practice (GCP) (EMA/CHMP/ICH/135/1995) 2016.
32. NIH , Data and Safety Monitoring Board (DSMB) Guidelines July 2018 https://www.nidcr.nih.gov/research/human-subjects-research/interventional-studies/data-and-safety-monitoring-board-guidelines
33. Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomized trials. BMJ 2010;340:c332..
34. Morris JA, Gardner MJ. Calculating confidence intervals for relative risks (odds ratios) and standardised ratios and rates. Stat Med 1988; 296: 1313-1316.
35. British Society of Audiology. Recommended Procedure: Pure-tone air conduction and bone-conduction threshold audiometry with and without masking (2011)
36. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999;8:135–60. 10.1191/096228099673819272.
37. Bearn DR, Sandy JR, Shaw WC. Photogrammetric Assessment of the Soft Tissue Profile in Unilateral Cleft Lip and Palate. Cleft Palate–Craniofacial Journal, November 2002, Vol. 39 No.
38. International Conference on Harmonisation. Topic E3 Structure and content of clinical study reports.(CPMP/ICH/137/95). 1995
39. International Conference on Harmonisation. Topic E9 Statistical Principles for Clinical Trials (CPMP/ICH/363/96). 1998.
40. Guidance on Statistical Analysis Plans.
41. Shuster JJ: Diagnostics for assumptions in moderate to large simple clinical trials: do they really help? Stat Med 2005, 24(16):2431–2438.
42. Campbell H, Dean CB. The consequences of proportional hazards based model selection. Statistics in Medicine. 2013
43. Raab GM, Day S, Sales J: How to select covariates to include in the analysis of a clinical trial. Control Clin Trials 2000, 21(4):330–342.
44. Freeman PR: The performance of the two-stage analysis of two treatment, two-period crossover trials. Stat Med 1989, 8(12):1421–1432.)
45. Kahan BC: Bias in randomised factorial trials. Stat Med 2013, 32(26):4540–4549.

23 Appendix

23.1 Definition of protocol deviations

Table 22: Definition of protocol deviations (Monitoring Plan v1.1 08Jul2013)

Protocol specification	Potential deviation(s)	Impact	Justification
Inclusion criteria			
Written informed proxy consent	Consent not obtained	**Major**	Violation of participant rights
One parent/carer a native language speaker	Neither parent or carer a native language speaker	**Major**	Statistical impact Primary outcome (insufficient velopharyngeal function) may be biased to poorer outcome in these patients. Primary analysis will be biased if violation is imbalanced between treatment groups. Safety impact None
Exclusion criteria			
Infants with syndromic cleft palate or severe developmental delay	Identification post randomisation	**Major**	Statistical impact Primary outcome (insufficient velopharyngeal function) may be biased to poorer outcome in these patients or may not be evaluable resulting in missing data. Primary analysis will be biased if occurrence of violation is imbalanced between treatment groups or a reduction in power caused by missing data. Safety impact None
Congenital sensorineural hearing loss or structural middle ear anomalies	Identification post randomisation	**Major**	Statistical impact Primary outcome (insufficient velopharyngeal function) may be biased to poorer outcome in these patients. Primary analysis will be biased if occurrence of violation is imbalanced between treatment groups. Safety impact None
One stage closure with the Sommerlad technique would be inappropriate	One stage closure with the Sommerlad technique not used at study surgery	**Major**	Statistical impact This will introduce greater variation in definition of the study treatment. This may bias the primary outcome. Unknown whether it will increase variation in the primary outcome. Safety impact No greater than standard medical care where the most appropriate surgical technique will be used.

Protocol specification	Potential deviation(s)	Impact	Justification
Submucous cleft palate	Submucous cleft palate is a different type of cleft palate condition than the study targeted isolated overt cleft palate condition	**Major**	Statistical impact Interpretations of results are intended only for isolated overt cleft palate medical condition. The inclusion of other types of cleft conditions in study summaries will add uncertainty in interpretation for the intended isolated overt cleft palate population. In addition, the presence of other cleft types may increase heterogeneity in the primary outcomes and therefore potential to decrease statistical power. Safety impact None
Treatment: timing of surgery			
Surgery at 6 months or 12 months should be within visit windows	Surgery outside visit window	**Major**	Statistical impact Patients would no longer be receiving the study intervention defined in protocol. Safety impact Major impact if a child was operated on earlier than assigned timing of surgery. In particular, surgery earlier than 6 months is expected to carry the highest risk of peri-operative complications. Minor or no impact if surgery was conducted later than assigned timing of surgery.
Withdrawal before surgery due to parent/guardian decision	(a) Refusal to keep to assigned timing of surgery (b) Withdrawal with no follow up	**Major**	Statistical impact (a) Bias to treatment comparison as choice of treatment is non-randomised (b) Missing data if not followed up Safety impact None
Study assessments			
Scheduled assessment visit at 30 days post-surgery: late complications	Missing or follow up assessment deviates from 30 day post-surgery date	**Minor**	Statistical impact No risk of bias to primary outcome (insufficient velopharyngeal function at age 3 and 5 years) Safety impact No impact. No visit window is specified in the protocol, but the number of day's deviation from 30 days will be described.

Protocol specification	Potential deviation(s)	Impact	Justification
Scheduled assessment visit at 12 months	Missing or visit occurs outside allowed window either side of visit date, or for 12-month patients occurs after surgery	**Minor**	Statistical impact No risk of bias to primary outcome (insufficient velopharyngeal function at age 3 and 5 years). However, there will be impact to secondary speech outcomes if speech visits are conducted in the 12 month group after the surgery has taken place the child is likely to be uncooperative or not perform as well as the surgery will be healing. Safety impact Minor for 6 month patients as surgery already performed. Minor for 12 month patients as timing of scheduled assessments are scheduled around the surgery date and not vice versa, in order to keep to the study timing of surgery.
Scheduled assessment visit at 3 years and 5 years	Missing or visit occurs outside allowed window either side of visit date.	**Major**	Statistical impact Data from both 3 and 5 years required for primary outcome. Loss of power and potential for bias. Safety impact Minor impact as surgery already performed.

23.2 Summary of outcomes

Table 23 provides a summary of each outcome, providing the time point of collection, source of outcome and relevant sections of the SAP. *Table 24-Table 27* provide the decoding of the columns in *Table 23.*

Table 23: Summary of TOPS outcomes

Outcome ID	Protocol ID	Outcome	Time point		Outcome				SAP sections	
			Visit (1)	Source (2)	Source (3)	(4)	Data access (5)	Access details (6)	Derivation	Analysis
Primary outcome										
PO_5y		Insufficient velopharyngeal function	5y	F13	2	5y_1 5y_3	2	CSV1 CSV2	18.6.1.2	18.6.1.3
Secondary outcomes										
	1	*Velopharyngeal function*								
SO_1a_5y	A	Velopharyngeal composite score summary (VPC-sum)	5y	F13	2	5y_1 5y_3	2	CSV1 CSV2	18.6.2.1.1	18.6.2.1.2
SO_1b_5y	B	Insufficient velopharyngeal function (VPC-rate)	5y	F13	2	5y_2	2	CSV3	18.6.2.1.1	18.6.2.1.2
	2	*Velopharyngeal function*								
SO_2a_3y	A	Insufficient velopharyngeal function (VPC-rate)	3y	F9	2	3y_2	2	CSV5	18.6.2.2.1	18.6.2.2.2
SO_2b_3y	B	Velopharyngeal insufficiency symptoms	3y	F9	2	3y_1	2	CSV4	18.6.2.2.1	18.6.2.2.2

Outcome ID	Protocol ID	Outcome	Time point		Outcome				SAP sections	
			Visit (1)	Source (2)	Source (3)	(4)	Data access (5)	Access details (6)	Derivation	Analysis
	3	*Canonical babbling*								
SO_3a_12m	A	Canonical babbling present	12m	F4	2	12m_1	2	CSV6	18.6.2.3.1	18.6.2.3.2
SO_3b_12m	B	Canonical babbling ratio	12m	F4	2	12m_1	2	CSV6	18.6.2.3.1	18.6.2.3.2
SO_3c_12m	C	Consonant inventory	12m	F4	2	12m_1	2	CSV6	18.6.2.3.1	18.6.2.3.2
	4	*Articulation*								
SO_4a_3y	A	Percent consonants correct (PCC)	3y	F9	2	3y_1	2	CSV4	18.6.2.4.1	18.6.2.4.2
SO_4b_3y	B	Percent correct placement (PCP)	3y	F9	2	3y_1	2	CSV4	18.6.2.4.1	18.6.2.4.2
SO_4c_3y	C	Percent correct manner (PCM)	3y	F9	2	3y_1	2	CSV4	18.6.2.4.1	18.6.2.4.2
SO_4d_3y	D	Non-oral consonant errors	3y	F9	2	3y_1	2	CSV4	18.6.2.4.1	18.6.2.4.2
SO_4e_3y	E	Oral consonant errors	3y	F9	2	3y_1	2	CSV4	18.6.2.4.1	18.6.2.4.2
	5	*Articulation*								
SO_5a_5y	A	Percent consonants correct (PCC)	5y	F13	2	5y_1	2	CSV1	18.6.2.5.1	18.6.2.5.2
SO_5b_5y	B	Percent correct placement (PCP)	5y	F13	2	5y_1	2	CSV1	18.6.2.5.1	18.6.2.5.2
SO_5c_5y	C	Percent correct manner (PCM)	5y	F13	2	5y_1	2	CSV1	18.6.2.5.1	18.6.2.5.2
SO_5d_5y	D	Non-oral consonant errors	5y	F13	2	5y_1	2	CSV1	18.6.2.5.1	18.6.2.5.2
SO_5e_5y	E	Oral consonant errors	5y	F13	2	5y_1	2	CSV1	18.6.2.5.1	18.6.2.5.2
	6	*Postoperative/long term complications*								
SO_6a_48h	A	Dehiscence	48h	F2	1	NA	1	F3/ F2	18.6.2.6.1	18.6.2.6.2

Outcome ID	Protocol ID	Outcome	Time point		Outcome				SAP sections	
			Visit (1)	Source (2)	Source (3)	(4)	Data access (5)	Access details (6)	Derivation	Analysis
SO_6a_30d	A	Dehiscence	30d	F3/F12	1	NA	1	F3	18.6.2.6.1	18.6.2.6.2
SO_6b_48h	B	Infection	48h	F2	1	NA	1	F3	18.6.2.6.1	18.6.2.6.2
SO_6b_30d	B	Infection	30d	F3	1	NA	1	F3	18.6.2.6.1	18.6.2.6.2
SO_6c_30d	C	Evidence of fistula	30d	F3/F12/F2	1	NA	1	F3	18.6.2.6.1	18.6.2.6.2
SO_6c_3y	C	Evidence of fistula	3y	F7	1	NA	1	F7	18.6.2.6.1	18.6.2.6.2
SO_6c_5y	C	Evidence of fistula	5y	F7	1	NA	1	F7	18.6.2.6.1	18.6.2.6.2
	7a	*Hearing level*								
SO_7ai_12m	I	Abnormal Transient Otoacoustic Emission (TEOAE)	12m	F6	1	NA	1	F6	18.6.2.7.1	18.6.2.7.2
SO_7aii_12m	Ii	Abnormal Sound field audiometry	12m	F6	1	NA	1	F6	18.6.2.7.1	18.6.2.7.2
	7b	*Hearing level*								
SO_7bi_3y	I	Abnormal Pure tone audiometry in at least one ear	3y	F8	1	NA	1	F8	18.6.2.7.1	18.6.2.7.2
SO_7bi_5y	I	Abnormal Pure tone audiometry in at least one ear	5y	F8	1	NA	1	F8	18.6.2.7.1	18.6.2.7.2
SO_7bii_3y	Ii	Abnormal Pure tone audiometry in both ears	3y	F8	1	NA	1	F8	18.6.2.7.1	18.6.2.7.2

Outcome ID	Protocol ID	Outcome	Time point		Outcome				SAP sections	
			Visit (1)	Source (2)	Source (3)	(4)	Data access (5)	Access details (6)	Derivation	Analysis
SO_7bii_5y	Ii	Abnormal Pure tone audiometry in both ears	5y	F8	1	NA	1	F8	18.6.2.7.1	18.6.2.7.2
SO_7biii_3y	Iii	Severity of better ear (normal, mild, moderate, severe, profound)	3y	F8	1	NA	1	F8	18.6.2.7.1	18.6.2.7.2
SO_7biii_5y	Iii	Severity of better ear (normal, mild, moderate, severe, profound)	5y	F8	1	NA	1	F8	18.6.2.7.1	18.6.2.7.2
	8	*Middle ear function*								
SO_8a_12m	A	Flat line Tympanogram in at least one ear	12m	F6	1	NA	1	F6	18.6.2.8.1	18.6.2.8.2
SO_8a_3y	A	Flat line Tympanogram in at least one ear	3y	F8	1	NA	1	F8	18.6.2.8.1	18.6.2.8.2
SO_8a_5y	A	Flat line Tympanogram in at least one ear	5y	F8	1	NA	1	F8	18.6.2.8.1	18.6.2.8.2
SO_8b_12m	B	Flat line Tympanogram in both ears	12m	F6	1	NA	1	F6	18.6.2.8.1	18.6.2.8.2
SO_8b_3y	B	Flat line Tympanogram in both ears	3y	F8	1	NA	1	F8	18.6.2.8.1	18.6.2.8.2
SO_8b_5y	B	Flat line Tympanogram in both ears	5y	F8	1	NA	1	F8	18.6.2.8.1	18.6.2.8.2
	9	*Dentofacial development*								
SO_9a_5y	A	Soft tissue ANB	5y	F16	3	NA	1	F16	18.6.2.9.1	18.6.2.9.2

Outcome ID	Protocol ID	Outcome	Time point		Outcome				SAP sections	
			Visit (1)	Source (2)	Source (3)	(4)	Data access (5)	Access details (6)	Derivation	Analysis
SO_9b_5y	B	Maxillary arch constriction score	5y	F17	4	NA	1	F17	18.6.2.9.1	18.6.2.9.2
	10	*Growth*								
SO_10a_12m	A	Nude weight	12m	F5	1	NA	1	F5	18.6.2.10.1	18.6.2.10.2
SO_10b_12m	B	Crown to heel length	12m	F5	1	NA	1	F5	18.6.2.10.1	18.6.2.10.2
SO_10c_12m	C	Occipitofrontal circumference	12m	F5	1	NA	1	F5	18.6.2.10.1	18.6.2.10.2

Note:
(1) 48h: 48 hours; 30d: 30 days; 12m: 12 months; 3y: 3 year; 5y: 5 year.
(2) Key provided in *Table 24*.
(3) Key provided in *Table 25*.
(4) Key provided in *Table 4*.
(5) Key provided in *Table 26*.
(6) Key provided in *Table 27*.

Table 24: CRF time variable

TS	Form	Item
F2	Form 2: Surgery and postoperative care	SURGERYD
F3	Form 3: Late Complications	VISDAT3D
F4	Form 4: 12 month speech assessment	VISDAT5D
F5	Form 5: 12 month growth	VISDAT6D
F6	Form 6: 12 month audiology	VISDT4D
F7	Form 7: Fistula assessment	VISDAT7D
F8	Form 8: Audiology data collection 3 and 5 years	VISDAT8D
F9	Form 9: 3 Year Speech Assessment	VISDA10D
F12	Form 12: Secondary Surgery	SECSRGD
F13	Form 13: 5 Year Speech Assessment	VISDA13D

Table 25: Outcome original source

OS	Original source
1	CRF
2	Speech recording
3	Photograph
4	Impression

Table 26: Data Access

DA	Data accessed
1	CRF
2	CSV

Table 27: Access details

AD	Access details
CSV1	*5 year speech CSV raw data and location – Phonetics [to be added when data available]*
CSV2	*5 year speech CSV raw data and location – Hypernasality [to be added when data available]*
CSV3	*5 year speech CSV raw data and location – Spontaneous Speech [to be added when data available]*
CSV4	TOPS\Randomisation\Speech meeting 2018-11\3 year data\ PhoneticsResults
CSV5	TOPS\Randomisation\Speech meeting 2018-11\3 year data\ VPCResults
CSV6	TOPS\Statistical Analysis\12 month speech outcomes\Raw data • raw1 summary_combined_25Oct2018.xlsx • DW012030_1year_a.mp4_EWWilladsen_DK.xlsx • DW012030_1year_a.mp4_EWWilladsen_DK_syllables.xlsx • DW012030_1year_b.mp4_EWWilladsen_DK.xlsx • DW012030_1year_b.mp4_EWWilladsen_DK_syllables.xlsx • FM211005 1 year audio back up file.wav_EWWilladsen_DK.xlsx • FM211005 1 year audio back up file.wav_EWWilladsen_DK_bak_syllables.xlsx • JC182001_1year_ a.wav_EWWilladsen_DK.xlsx • JC182001_1year_ a.wav_EWWilladsen_DK_syllables.xlsx • JC182001_1year_ b.wav_EWWilladsen_DK.xlsx • JC182001_1year_ b.wav_EWWilladsen_DK_syllables.xlsx • MB142028_1year_a.mp4_EWWilladsen_DK.xlsx

AD	Access details
	• MB142028_1year_a.mp4_EWWilladsen_DK_bak_syllables.xlsx • MB142028_1year_b.mp4_EWwilladsen_DK.xlsx • MB142028_1year_b.mp4_EWwilladsen_DK_syllables.xlsx
F2	Form 2: Surgery and Postoperative care
F3	Form 3: Late Complications
F4	Form 4: 12 month speech assessment
F5	Form 5: 12 month growth
F6	Form 6: 12 month audiology
F7	Form 7: Fistula Assessment
F8	Form 8: Audiology data collection 3 and 5 years
F9	Form 9: 3 Year Speech Assessment
F12	Form 12: Secondary Surgery
F13	Form 13: 5 Year Speech Assessment
F16	Form 16: Soft Tissue ANB
F17	Form 17: Maxillary Arch Constriction Score

23.3 TimeStamper example output

Table 28: 12 months example TimeStamper output

File	SLT	Created	Stavelse	Stavelse Count	Kanonisk Stavelse	Kanonisk Stavelse Count	Total	File	SLT	Created	Comment	Sounds		
XY000000_12months_A.mp4	abaa.uk	10/11/2012 01:23:45	0.827586	96	0.172414	20	116		XY000000_12months_A.mp4	abaa.uk	10/11/2012 01:23:45	canonical	ma	na
XY000000_12months_B.mp4	abaa.uk	10/11/2012 01:25:45	0.307692	24	0.692308	54	78		XY000000_12months_B.mp4	abaa.uk	10/11/2012 01:25:45	Not canonical	ba	
XY000000_12months_B.mp4	bcaa.uk	11/11/2012 01:23:45	0.360759	57	0.639241	101	158		XY000000_12months_B.mp4	bcaa.uk	11/11/2012 01:23:45	canonical	la	ha

23.4 SPEAK example output

23.4.1 3 year

Table 29: 3 year example SPEAK output for Phonetics

FileName	Modified Date	Assessor Name	Assessor Init	SURGINIT	Site Code	PatientId	Randomisation Number	Test Order	DataFile	Language	Word Idx	Notes	Sounds Idx	Correct	Quality
E:\Data\...\..Phonetics Transcription.json	10/11/2012 01:23:45	A.Baa	AB	XY	00	0000	XY000000	01		English	01			1	0
E:\Data\...\..Phonetics Transcription.json	10/11/2012 01:23:45	A.Baa	AB	XY	00	0000	XY000000	01		English	02		24	0	0
E:\Data\...\...Phonetics Transcription.json	10/11/2012 01:23:45	A.Baa	AB	XY	00	0000	XY000000	01		English	03	File Missing		0	0
E:\Data\...\...Phonetics Transcription.json	10/11/2012 01:23:45	A.Baa	AB	XY	00	0000	XY000000	01		English	01	Could Not Assess		0	0

Table 30: 3 year example SPEAK output for VPC-rate

FileName	Date Modified	Assessor	DataFile	InDataDirectory	Language	OutDataDirectory	OutDataFileName	Step1	Step2	Comment
E:\Data\...\VPC-Ess…_3y.mp4_.json	10/11/2012 01:23:45	Baa_A	Ess_XY000000_3y.mp4	C:\...\..xy000000\VPC Analysis	.	C:\Users\tops\SPEAK-Data	VPC Ess.... XY000000_3y.mp4_.json	Yes		
E:\Data\...\VPC-Ess…_3y.mp4_.json	10/11/2012 01:23:45	Baa_A	Ess_XY000000_3y.mp4	C:\...\..xy000000\VPC Analysis	.	C:\Users\tops\SPEAK-Data	VPC Ess.... XY000000_3y.mp4_.json	No	Marginal Incompetent	
E:\Data\...\VPC-Ess…_3y.mp4_.json	10/11/2012 01:23:45	Baa_A	Ess_XY000000_3y.mp4	C:\...\..xy000000\VPC Analysis	.	C:\Users\tops\SPEAK-Data	VPC Ess.... XY000000_3y.mp4_.json	No	Incompetent	

FileName	Date Modified	Assessor	DataFile	InDataDirectory	Language	OutDataDirectory	OutDataFileName	Step1	Step2	Comment
E:\Data\...\VPC-Ess…_3y.mp4_.json	10/11/2012 01:23:45	Baa_A	Ess_XY000000_3y.mp4	C:\...\..xy000000\VPC Analysis	.	C:\Users\tops\SPEAK-Data	VPC Ess....XY000000_3y.mp4_.json	Could Not Assess		Reason

23.4.2 5 year

Table 31: 5 year example SPEAK output for Phonetics

SltNumber	Language	PatientId	OrdeNo	AssessmntId	AssessmentStatus	StartTime	EndTime	MediaFile	WordId	Word	TargetSounds	TargetSoundsIdx	Word AssessmentStatus	ActualSounds	ActualSoundsIdx
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Completed	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK..._3y_puppy.mp4	1	pea	p^h	65	Correct		
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Completed	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK..._3y_puppy.mp4	1	pea	p^h	65	Incorrect	b	2
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Completed	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK..._3y_puppy.mp4	1	pea	p^h	65	MediaNotAvailable		
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Completed	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK..._3y_puppy.mp4	1	pea	p^h	65	CouldNotAssess		
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Cancelled	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK..._3y_puppy.mp4	1	pea	p^h	65	Correct		

Table 32: 5 year example SPEAK output for Hypernasality

SltNumber	Language	PatientId	OrderNo	AssessmentId	AssessmentStatus	StartTime	EndTime	MediaFile	WithinNormalLimits	DegreeOfSeverity	CouldNotAssessReason	OtherText
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Completed	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK…_3y.mp4	Yes			
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Completed	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK…_3y.mp4	No	Mild		
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Completed	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK…_3y.mp4	No	Moderate		
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Completed	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK…_3y.mp4	No	Severe		
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Completed	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK…_3y.mp4	Could not Assess		Other	Reason

Table 33: 5 year example SPEAK output for VPC-rate

SltNumber	Language	PatientId	Order No	AssessmentId	Assessment Status	StartTime	EndTime	MediaFile	WithinNormalLimits	DegreeOfSeverity	CouldNotAssessReason	OtherText
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Completed	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK…_3y.mp4	Yes			
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Completed	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK…_3y.mp4	No	Marginal Incompetent		
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Completed	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK…_3y.mp4	No	Incompetent		
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Completed	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK…_3y.mp4	Could not Assess		Not enough speech	
1	English	XY000000	1	09aksm91-sos39dj-29jdp	Completed	10/11/2012 01:23:45	10/11/2012 01:34:56	C:\SPEAK…_3y.mp4	Could Not assess		Other	Reason

23.5 Mapping structures

Table 34: 12 month speech mapping structure

Preferred Notation	Oral Stops (1)	Equivalent	Similar	Type of sound
gla	g	gla		Consonant cluster
dla	d	dla		Consonant cluster
tʃa		tʃa / tsa / ʦa / ʧa		Voiceless postalveolar affricate
ŋa		ŋa / nga/ ŋga / aŋ /ang	ɲa/ ʁa/ Na	Velar nasal
ba	b	ba/ bwa / ab	ʙa/Ba	Voiced bilabial plosive
ma		ma/ mwa/ mja/am / me / ma/ Ma		Bibalbial nasal
na		na/nan/an/ne/ nʔa	nja	Alveolar nasal
da	d	da/ / Da	dja/ɟa	Voiced dental/ alveolar plosive
bla	b	bla		Consonant cluster
ɬa		/ ɬa/ aɬ	ʓa	Unvoiced alveolar lateral fricative
ga	g	ga/	ɢa/ gʁ/ Ga	Voiced velar plosive
wa		Wa/wa/we	βa / va/ ʋa/ ʊa	Voiced velar-labial
ra		ra	ɹa	Voiced velar-labial approximant
ja		ja/ ya/ aj/ ai/ Ja/ ʔja	ɯa	Voiced palatal approximant
pa	p	pa		Unvoiced bilabial plosive
ha		ha		unvoiced glottal fricative
la		la	ɭa/ ʎa	Voiced alveolar lateral approximant
ka	k	ka	qa	Unvoiced velar plosive
ʔa		ʔa/ aʔ/ ʔ		Unvoiced glottal plosive
sa		sa	ça/ Ca/ ɕa/ θa/ za/ ʃa/ ʂa/ ʒa	Unvoiced alveolar fricative
ta	t	ta/te/ at/ca		Unvoiced dental/ alveolar plosive
ða	d	ða/ að		Voiced dental fricative
xa		Xa/ ʁa/ ɣa/ ʀa/ xa/χa		Unvoiced uvular fricative
ʕa		aʕ/ ʕa		Unvoiced pharyngeal fricative
fa		fa	ɸa	Unvoiced labiodental fricative
ħa		ħa		Unvoiced pharyngeal fricative

Note:
(1) Oral Stops are not linked to outcome, but included for assistance in any post-hoc analysis.

Table 35: 3 year phonetics mapping structure

Sound	Idx	Manner of articulation	Voicing	Place of articulation	Text to appear in SPEAK	Correct production if:	VPI symptoms present if:	Oral consonant errors if:	Not oral consonant error if:	Non-oral consonant errors if:	Not non-oral consonant error if:
p	1	Plosive	Unvoiced	Bilabial	Unvoiced bilabial plosive	n.a.	p̊, p/ ̊, p/˜	n.a.	n.a	n.a	n.a
b	2	Plosive	Voiced	Bilabial	Voiced bilabial plosive	n.a.	b̥, b̃, b/ ̊, b/ ̃	n.a.	n.a	n.a	n.a
t	3	Plosive	Unvoiced	Dental/Alveolar	Unvoiced dental/alveolar plosive	n.a.	t̥, t/ ̊, t/˜	n.a.	n.a	n.a	n.a
d	4	Plosive	Voiced	Dental/Alveolar	Voiced dental/alveolar plosive	n.a.	d̥, d̃, d/ ̊, d/ ̃	n.a.	n.a	n.a	n.a
ʈ	5	Plosive	Unvoiced	Retroflex	Unvoiced retroflex plosive	n.a.	ʈ/ ̊, ʈ/˜	n.a.	n.a	n.a	n.a
ɖ	6	Plosive	Voiced	Retroflex	Voiced retroflex plosive	n.a.	ɖ/ ̊, ɖ/˜	n.a.	n.a	n.a	n.a
c	7	Plosive	Unvoiced	Palatal	Unvoiced palatal plosive	n.a.	c/ ̊, c/˜	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a	n.a
ɟ	8	Plosive	Voiced	Palatal	Voiced palatal plosive	n.a.	ɟ/ ̊, ɟ/˜	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a	n.a
k	9	Plosive	Unvoiced	Velar	Unvoiced velar plosive	n.a.	k̥, k/ ̊, k/˜	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a	n.a
g	10	Plosive	Voiced	Velar	Voiced velar plosive	n.a.	g̊, g̃, g̰/ ̊, g/	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a	n.a
q	11	Plosive	Unvoiced	Uvular	Unvoiced uvular plosive	n.a.	q/ ̊, q/˜	Always oral consonant error	Always oral consonant error	n.a	n.a
ɢ	12	Plosive	Voiced	Uvular	Voiced uvular plosive	n.a.	ɢ/ ̊, ɢ/˜	Always oral consonant error	Always oral consonant error	n.a	n.a
ʔ[(1)]	13[(1)]	Plosive	Unvoiced	Glottal	Glottal plosive	n.a.	n.a.	n.a	n.a	always non-oral consonant error	always non-oral consonant error
m[(2)]	14[(2)]	Nasal	Voiced	Bilabial	Bilabial nasal	n.a.	b, v, ʊ	n.a.	n.a	p, d, t, g, k, s, f, l	n.a

Sound	Idx	Manner of articulation	Voicing	Place of articulation	Text to appear in SPEAK	Correct production if:	VPI symptoms present if:	Oral consonant errors if:	Not oral consonant error if:	Non-oral consonant errors if:	Not non-oral consonant error if:
ɱ(2)	15(2)	Nasal	Voiced	Labiodental	Labiodental nasal	n.a.	b, v, ʊ	n.a.	n.a	p, d, t, g, k, s, f, l	n.a
n(2)	16(2)	Nasal	Voiced	Alveolar	Alveolar nasal	n	d, l	n.a.	n.a	p, b, t, g, k, f, s, v, ʊ	n.a
ɳ(2)	17(2)	Nasal	Voiced	Retroflex	Retroflex nasal	n.a.	d	n.a.	n.a	p, b, t, g, k, f, s, v, ʊ, l	n.a
ɲ	18	Nasal	Voiced	Palatal	Palatal nasal	n.a.	n.a.	n	n.a	p, b, t, d, k, g, f, v, ʊ, s, l	n.a
ŋ (2)	19(2)	Nasal	Voiced	Velar	Velar nasal	n.a.	g	n	n.a	p, b, t, d, k, f, v, ʊ ,s, l	n.a
ɴ	20	Nasal	Voiced	Uvular	Uvular nasal	n.a.	n.a.	n	n.a	p, b, t, d, k, g, f, v, ʊ, s, l	n.a
ʙ	21	Trill	Voiced	Bilabial	Voiced bilabial trill	n.a.	ʙ/ ̊, ʙ/˜	n.a.	n.a	n.a	n.a
r	22	Trill	Voiced	Dental/Alveolar	Voiced dental/alveolar trill	n.a.	r/ ̊, r/˜	n.a.	n.a	n.a	n.a
ʀ	23	Trill	Voiced	Uvular	Voiced uvular trill	n.a.	ʀ/ ̊, ʀ/˜	Always oral consonant error	Always oral consonant error	n.a	n.a
ⱱ	24	Flap	Voiced	Labiodental	Voiced labiodental flap	n.a.	ⱱ/ ̊, ⱱ/˜	n.a.	n.a	n.a	n.a
ɾ	25	Flap or tap	Voiced	Alveolar	Voiced alveolar flap/tap	n.a.	ɾ/ ̊, ɾ/˜	n.a.	n.a	n.a	n.a
ɽ	26	Flap or tap	Voiced	Retroflex	Voiced retroflex flap/tap	n.a.	ɽ/ ̊, ɽ/˜	n.a.	n.a.	n.a	n.a
ɸ	27	Fricative	Unvoiced	Bilabial	Unvoiced bilabial fricative	n.a.	ɸ/ ̊, ɸ/˜	n.a.	n.a	n.a.	n.a.
β	28	Fricative	Voiced	Bilabial	Voiced bilabial fricative	n.a.	β/ ̊, β/˜	n.a.	n.a	n.a.	n.a.
f	29	Fricative	Unvoiced	Labiodental	Unvoiced labiodental fricative	n.a.	f̊, f/ ̊, f/˜	n.a.	n.a	n.a.	n.a.

Sound	Idx	Manner of articulation	Voicing	Place of articulation	Text to appear in SPEAK	Correct production if:	VPI symptoms present if:	Oral consonant errors if:	Not oral consonant error if:	Non-oral consonant errors if:	Not non-oral consonant error if:
v	30	Fricative	Voiced	Labiodental	Voiced labiodental fricative	n.a.	ʋ̃, ʋ/̊, ʋ/̃	n.a.	n.a	n.a.	n.a.
θ	31	Fricative	Unvoiced	Dental	Unvoiced dental fricative	n.a.	θ/̊, θ/̃	n.a.	n.a	n.a.	n.a.
ð	32	Fricative	Voiced	Dental	Voiced dental fricative	n.a.	ð/̊, ð/̃	n.a.	n.a	n.a.	n.a.
s	33	Fricative	Unvoiced	Alveolar	Unvoiced alveolar fricative	n.a.	s̊, s/̊, s/̃	n.a.	n.a	n.a.	n.a.
z	34	Fricative	Voiced	Alveolar	Voiced alveolar fricative	n.a.	z/̊, z/̃	n.a.	n.a	n.a.	n.a.
ʃ	35	Fricative	Unvoiced	Postalveolar	Unvoiced postalveolar fricative	n.a.	ʃ/̊, ʃ/̃	n.a.	n.a	n.a.	n.a.
ʒ	36	Fricative	Voiced	Postalveolar	Voiced postalveolar fricative	n.a.	ʒ/̊, ʒ/̃	n.a.	n.a	n.a.	n.a.
ʂ	37	Fricative	Unvoiced	Retroflex	Unvoiced retroflex fricative	n.a.	ʂ/̊, ʂ/̃	n.a.	n.a	n.a.	n.a.
ʐ	38	Fricative	Voiced	Retroflex	Voiced retroflex fricative	n.a.	ʐ/̊, ʐ/̃	n.a.	n.a	n.a.	n.a.
ɕ	39	Fricative	Unvoiced	Alveolo-palatal	Unvoiced alveolo-palatal fricative	n.a.	ɕ/̊, ɕ/̃	n.a.	n.a	n.a.	n.a.
ʑ	40	Fricative	Voiced	Alveolo-palatal	Voiced alveolo-palatal fricative	n.a.	ʑ/̊, ʑ/̃	n.a.	n.a	n.a.	n.a.
ç	41	Fricative	Unvoiced	Palatal	Unvoiced palatal fricative	n.a.	ç/̊, ç/̃	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a.	n.a.
ʝ	42	Fricative	Voiced	Palatal	Voiced palatal fricative	n.a.	ʝ/̊, ʝ/̃	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a.	n.a.
x	43	Fricative	Unvoiced	Velar	Unvoiced velar fricative	n.a.	x/̊, x/̃	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a.	n.a.
ɣ	44	Fricative	Voiced	Velar	Voiced velar fricative	n.a.	ɣ/̊, ɣ/̃	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a.	n.a.

Sound	Idx	Manner of articulation	Voicing	Place of articulation	Text to appear in SPEAK	Correct production if:	VPI symptoms present if:	Oral consonant errors if:	Not oral consonant error if:	Non-oral consonant errors if:	Not non-oral consonant error if:
χ	45	Fricative	Unvoiced	Uvular	Unvoiced uvular fricative	n.a.	χ/ ̊, χ/˜	Always oral consonant error	Always oral consonant error	n.a.	n.a.
ʁ	46	Fricative	Voiced	Uvular	Voiced uvular fricative	n.a.	ʁ/ ̊, ʁ/˜	Always oral consonant error	Always oral consonant error	n.a.	n.a.
ħ	47	Fricative	Unvoiced	Pharyngeal	Unvoiced pharyngeal fricative	n.a.	n.a.	n.a.	n.a.	always non-oral consonant error	always non-oral consonant error
ʕ	48	Fricative	Voiced	Pharyngeal	Voiced pharyngeal fricative	n.a.	n.a.	n.a.	n.a.	always non-oral consonant error	always non-oral consonant error
h	49	Fricative	Unvoiced	Glottal	Unvoiced glottal fricative	n.a.	n.a.	n.a.	n.a.	always non-oral consonant error	always non-oral consonant error
ɦ	50	Fricative	Voiced	Glottal	Voiced glottal fricative	n.a.	n.a.	n.a.	n.a.	always non-oral consonant error	always non-oral consonant error
ɬ	51	Lateral, fricative	Unvoiced	Alveolar	Unvoiced alveolar lateral fricative	n.a.	ɬ/ ̊, ɬ/˜	n.a.	n.a.	n.a.	n.a.
ɮ	52	Lateral, fricative	Voiced	Alveolar	Voiced alveolar lateral fricative	n.a.	ɮ/ ̊, ɮ/˜	n.a.	n.a.	n.a.	n.a.
ʋ	53	Approximant	Voiced	Labiodental	Voiced labiodental approximant	n.a.	ʋ̃, ʋ/ ̊, ʋ/˜	n.a.	n.a.	n.a.	n.a.
ɹ	54	Approximant	Voiced	Alveolar	Voiced alveolar approximant	n.a.	ɹ/ ̊, ɹ/˜	n.a.	n.a.	n.a.	n.a.
ɻ	55	Approximant	Voiced	Retroflex	Voiced retroflex approximant	n.a.	ɻ/ ̊, ɻ/˜	n.a.	n.a.	n.a.	n.a.
j	56	Approximant	Voiced	Palatal	Voiced palatal approximant	n.a.	j/ ̊, j/˜	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.

Sound	Idx	Manner of articulation	Voicing	Place of articulation	Text to appear in SPEAK	Correct production if:	VPI symptoms present if:	Oral consonant errors if:	Not oral consonant error if:	Non-oral consonant errors if:	Not non-oral consonant error if:
ɯ̞	57	Approximant	Voiced	Velar	Voiced velar approximant	n.a.	ɯ̞/ ̊, ɯ̞/ ̃	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.
l	58	Lateral approximant	Voiced	Alveolar	Voiced alveolar lateral approximant	n.a.	l/ ̊, l/ ̃	n.a.	n.a.	n.a.	n.a.
ɭ	59	Lateral approximant	Voiced	Retroflex	Voiced retroflex lateral approximant	n.a.	ɭ/ ̊, ɭ/ ̃	n.a.	n.a.	n.a.	n.a.
ʎ	60	Lateral approximant	Voiced	Palatal	Voiced palatal lateral approximant	n.a.	ʎ/ ̊, ʎ/ ̃	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.
ʟ	61	Lateral approximant	Voiced	Velar	Voiced velar lateral approximant	n.a.	ʟ/ ̊, ʟ/ ̃	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.
w	62	Approximant	Voiced	Velar labial	Voiced velar-labial approximant	n.a.	w/ ̊, w/ ̃	n.a.	n.a.	n.a.	n.a.
ɧ	63	Fricative	Unvoiced	Postalveolar and velar	Unvoiced postalveolar-velar fricative	n.a.	ɧ/ ̊, ɧ/ ̃	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.
Ø	64	Omission of sound	n.a.	n.a.	Sound omission	n.a.	n.a.	n.a.	n.a.	always non-oral consonant error	always non-oral consonant error
pʰ	65	Plosive	Unvoiced	Bilabial	Aspirated unvoiced bilabial plosive	n.a.	pʰ/ ̊, pʰ/ ̃	n.a.	n.a.	n.a.	n.a.
tʰ	66	Plosive	Unvoiced	Alveolar	Aspirated unvoiced alveolar plosive	n.a.	tʰ/ ̊, tʰ/ ̃	n.a.	n.a.	n.a.	n.a.
tˢ	68	Plosive	Unvoiced	Alveolar	Affricated unvoiced alveolar plosive	n.a.	tˢ/ ̊, tˢ/ ̃	n.a.	n.a.	n.a.	n.a.
kʰ	67	Plosive	Unvoiced	Velar	Aspirated unvoiced velar plosive	n.a.	kʰ/ ̊, kʰ/ ̃	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.

Sound	Idx	Manner of articulation	Voicing	Place of articulation	Text to appear in SPEAK	Correct production if:	VPI symptoms present if:	Oral consonant errors if:	Not oral consonant error if:	Non-oral consonant errors if:	Not non-oral consonant error if:
b̥́	69	Plosive	Voiced	Bilabial	Nasal emission unvoiced bilabial plosive	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
d̥́	70	Plosive	Voiced	Alveolar	Nasal emission unvoiced alveolar plosive	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
g̊̃	71	Plosive	Voiced	Velar	Nasal emission unvoiced velar plosive	n.a.	Always VPI-symptom	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.
p̊̃	72	Plosive	Unvoiced	Bilabial	Nasal emission unvoiced bilabial plosive	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
t́	73	Plosive	Unvoiced	Alveolar	Nasal emission unvoiced alveolar plosive	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
ḱ	74	Plosive	Unvoiced	Velar	Nasal emission unvoiced velar plosive	n.a.	Always VPI-symptom	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.
f́	75	Fricative	Unvoiced	Labiodental	Nasal emission unvoiced labiodental fricative	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
s̊̃	76	Fricative	Unvoiced	Alveolar	Nasal emission unvoiced alveolar fricative	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
b̃	85	Plosive	Voiced	Bilabial	Nasalised voiced bilabial plosive	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
d̃	86	Plosive	Voiced	Alveolar	Nasalised voiced alveolar plosive	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
g̃	87	Plosive	Voiced	Velar	Nasalised voiced velar plosive	n.a.	Always VPI-symptom	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.

Sound	Idx	Manner of articulation	Voicing	Place of articulation	Text to appear in SPEAK	Correct production if:	VPI symptoms present if:	Oral consonant errors if:	Not oral consonant error if:	Non-oral consonant errors if:	Not non-oral consonant error if:
ʊ̃	88	Approximant	Voiced	Labiodental	Nasalised voiced labiodental approximant	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
n̪[(2)]	77[(2)]	Nasal	Voiced	Interdental	Interdental nasal	n.a.	d	n.a.	n.a.	p, b, t, g, k, f, s, v, ʊl	n.a.
t̪	79	Plosive	Unvoiced	Interdental	Interdental unvoiced plosive	n.a.	t̪/ ˟, t̪/˜	n.a.	n.a.	n.a.	n.a.
d̪	78	Plosive	Voiced	Interdental	Interdental voiced plosive	n.a.	d̪/ ˟, d̪/˜	n.a.	n.a.	n.a.	n.a.
θ̪	80	Fricative	Unvoiced	Interdental	Interdental unvoiced fricative	n.a.	θ̪/ ˟, θ̪˜	n.a.	n.a.	n.a.	n.a.
m̥̃	81	Active nasal fricative	Unvoiced	Bilabial	Bilabial active nasal fricative	n.a.	n.a.	n.a	n.a	always non-oral consonant error	always non-oral consonant error
ñ̥	82	Active nasal fricative	Unvoiced	Alveolar	Alveolar active nasal fricative	n.a.	n.a.	n.a	n.a.	always non-oral consonant error	always non-oral consonant error
ŋ̥̃	83	Active nasal fricative	Unvoiced	Velar	Velar active nasal fricative	n.a.	n.a.	n.a	n.a.	always non-oral consonant error	always non-oral consonant error
Ń̥	84	Active nasal fricative	Unvoiced	Unknown	Active nasal fricative, place unspecified	n.a.	n.a.	n.a	n.a	always non-oral consonant error	always non-oral consonant error
ʰ	89	Aspiration	n.a.	n.a.	Aspiration	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
˜	92	Nasalised	n.a.	n.a.	Nasalisation	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
˟	90	Nasal emission	n.a.	n.a.	Nasal emission	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
˷	91	Denasalised	n.a.	n.a.	Denasalisation	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Notes:
(1) Any sound in combination with a glottal (Idx 13) will be treated as a glottal only (Idx 13).
(2) Exceptions to the VPI symptoms interpretation (IDx: 14,15,16,17,19,77) which treats the Sounds and VPI symptoms present column the opposite way around, see *Section 18.5.2.1.2.*

Table 36: 5 year phonetics mapping structure

Sound	Idx	Manner of articulation	Voicing	Place of articulation	Text to appear in SPEAK	Correct production if:	VPI symptoms present if:	Oral consonant error if:	Not oral consonant error if:	Non-oral consonant error if:	Not non-oral consonant error if:
p	1	Plosive	Unvoiced	Bilabial	Unvoiced bilabial plosive	n.a.	p̰, p/ ̞, p/ ̃, p/¬	n.a.	n.a	n.a	n.a
b	2	Plosive	Voiced	Bilabial	Voiced bilabial plosive	n.a.	b̰, b̃, b/ ̞, b/ ̃, b/¬	n.a.	n.a	n.a	n.a
t	3	Plosive	Unvoiced	Dental/Alveolar	Unvoiced dental/alveolar plosive	n.a.	t̰, t/ ̞, t/̃ ,t/¬	n.a.	n.a	n.a	n.a
d	4	Plosive	Voiced	Dental/Alveolar	Voiced dental/alveolar plosive	n.a.	d̰, d̃, d/ ̞, d/ ̃, d/¬	n.a.	n.a	n.a	n.a
ʈ	5	Plosive	Unvoiced	Retroflex	Unvoiced retroflex plosive	n.a.	ʈ/ ̞, ʈ/ ̃, ʈ/¬	n.a.	n.a	n.a	n.a
ɖ	6	Plosive	Voiced	Retroflex	Voiced retroflex plosive	n.a.	ɖ/ ̞, ɖ/ ̃, ɖ/¬	n.a.	n.a	n.a	n.a
c	7	Plosive	Unvoiced	Palatal	Unvoiced palatal plosive	n.a.	c/ ̞, c/ ̃, c/¬	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a	n.a
ɟ	8	Plosive	Voiced	Palatal	Voiced palatal plosive	n.a.	ɟ/ ̞, ɟ/ ̃, ɟ/¬	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a	n.a
k	9	Plosive	Unvoiced	Velar	Unvoiced velar plosive	n.a.	k̰, k/ ̞, k/ ̃, k/¬	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a	n.a
g	10	Plosive	Voiced	Velar	Voiced velar plosive	n.a.	g̰, g̃, g/ ̞, g/ ̃, g/¬	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a	n.a
q	11	Plosive	Unvoiced	Uvular	Unvoiced uvular plosive	n.a.	q/ ̞, q/ ̃, q/¬	Always oral consonant error	Always oral consonant error	n.a	n.a
ɢ	12	Plosive	Voiced	Uvular	Voiced uvular plosive	n.a.	ɢ/ ̞, ɢ/ ̃, ɢ/¬	Always oral consonant error	Always oral consonant error	n.a	n.a
ʔ[(1)]	13[(1)]	Plosive	Unvoiced	Glottal	Glottal plosive	n.a.	n.a.	n.a	n.a	always non-oral consonant error	always non-oral consonant error
m[(2)]	14[(2)]	Nasal	Voiced	Bilabial	Bilabial nasal	n.a.	b, v, ʊ	n.a.	n.a	p, d, t, g, k, s, f, l	n.a
ɱ[(2)]	15[(2)]	Nasal	Voiced	Labiodental	Labiodental nasal	n.a.	b, v, ʊ	n.a.	n.a	p, d, t, g, k, s, f, l	n.a

Sound	Idx	Manner of articulation	Voicing	Place of articulation	Text to appear in SPEAK	Correct production if:	VPI symptoms present if:	Oral consonant error if:	Not oral consonant error if:	Non-oral consonant error if:	Not non-oral consonant error if:
ɳ[(2)]	16[(2)]	Nasal	Voiced	Alveolar	Alveolar nasal	n.a.	d, l	n.a.	n.a	p, b, t, g, k, f, s, v, ʊ	n.a
ɳ[(2)]	17[(2)]	Nasal	Voiced	Retroflex	Retroflex nasal	n.a.	d	n.a.	n.a	p, b, t, g, k, f, s, v, ʊ, l	n.a
ɲ	18	Nasal	Voiced	Palatal	Palatal nasal	n.a.	n.a.	n	n.a.	p, b, t, d, g, k, f, v, ʊ, s, l	n.a.
ŋ[(2)]	19[(2)]	Nasal	Voiced	Velar	Velar nasal	n.a.	g	n	n.a.	p, b, t, d, k, f, v, ʊ, s, l	n.a.
ɴ	20	Nasal	Voiced	Uvular	Uvular nasal	n.a.	n.a.	n	n.a.	p, b, t, d, g, k, f, v, ʊ, s, l	n.a.
ʙ	21	Trill	Voiced	Bilabial	Voiced bilabial trill	n.a.	ʙ/ ̊, ʙ/˜	n.a.	n.a	n.a	n.a
r	22	Trill	Voiced	Dental/Alveolar	Voiced dental/alveolar trill	n.a.	r/ ̊, r/˜	n.a.	n.a	n.a	n.a
ʀ	23	Trill	Voiced	Uvular	Voiced uvular trill	n.a.	ʀ/ ̊, ʀ/˜	Always oral consonant error	Always oral consonant error	n.a	n.a
ⱱ	24	Flap	Voiced	Labiodental	Voiced labiodental flap	n.a.	ⱱ/ ̊, ⱱ/˜	n.a.	n.a	n.a	n.a
ɾ	25	Flap or tap	Voiced	Alveolar	Voiced alveolar flap/tap	n.a.	ɾ/ ̊, ɾ/˜	n.a.	n.a	n.a	n.a
ɽ	26	Flap or tap	Voiced	Retroflex	Voiced retroflex flap/tap	n.a.	ɽ/ ̊, ɽ/˜	n.a.	n.a.	n.a	n.a
ɸ	27	Fricative	Unvoiced	Bilabial	Unvoiced bilabial fricative	n.a.	ɸ/ ̊, ɸ/˜, ɸ/¬	n.a.	n.a	n.a.	n.a.
β	28	Fricative	Voiced	Bilabial	Voiced bilabial fricative	n.a.	β/ ̊, β/˜, β/¬	n.a.	n.a	n.a.	n.a.
f	29	Fricative	Unvoiced	Labiodental	Unvoiced labiodental fricative	n.a.	f̊, f/ ̊, f/˜,f/¬	n.a.	n.a	n.a.	n.a.
v	30	Fricative	Voiced	Labiodental	Voiced labiodental fricative	n.a.	v/ ̊, v/˜, v/¬, ṽ	n.a.	n.a	n.a.	n.a.

Sound	Idx	Manner of articulation	Voicing	Place of articulation	Text to appear in SPEAK	Correct production if:	VPI symptoms present if:	Oral consonant error if:	Not oral consonant error if:	Non-oral consonant error if:	Not non-oral consonant error if:
θ	31	Fricative	Unvoiced	Dental	Unvoiced dental fricative	n.a.	θ/ ̊, θ/ ̃, θ/¬	n.a.	n.a	n.a.	n.a.
ð	32	Fricative	Voiced	Dental	Voiced dental fricative	n.a.	ð/ ̊, ð/ ̃, ð/¬	n.a.	n.a	n.a.	n.a.
s	33	Fricative	Unvoiced	Alveolar	Unvoiced alveolar fricative	n.a.	s̊, s/ ̊, s/ ̃, s/¬	n.a.	n.a	n.a.	n.a.
z	34	Fricative	Voiced	Alveolar	Voiced alveolar fricative	n.a.	z/ ̊, z/ ̃, z/¬	n.a.	n.a	n.a.	n.a.
ʃ	35	Fricative	Unvoiced	Postalveolar	Unvoiced postalveolar fricative	n.a.	ʃ/ ̊, ʃ/ ̃, ʃ/¬	n.a.	n.a	n.a.	n.a.
ʒ	36	Fricative	Voiced	Postalveolar	Voiced postalveolar fricative	n.a.	ʒ/ ̊, ʒ/ ̃, ʒ/¬	n.a.	n.a	n.a.	n.a.
ʂ	37	Fricative	Unvoiced	Retroflex	Unvoiced retroflex fricative	n.a.	ʂ/ ̊, ʂ/ ̃, ʂ/¬	n.a.	n.a	n.a.	n.a.
ʐ	38	Fricative	Voiced	Retroflex	Voiced retroflex fricative	n.a.	ʐ/ ̊, ʐ/ ̃, ʐ/¬	n.a.	n.a	n.a.	n.a.
ɕ	39	Fricative	Unvoiced	Alveolo-palatal	Unvoiced alveolo-palatal fricative	n.a.	ɕ/ ̊, ɕ/ ̃, ɕ/¬	n.a.	n.a	n.a.	n.a.
ʑ	40	Fricative	Voiced	Alveolo-palatal	Voiced alveolo-palatal fricative	n.a.	ʑ/ ̊, ʑ/ ̃, ʑ/¬	n.a.	n.a	n.a.	n.a.
ç	41	Fricative	Unvoiced	Palatal	Unvoiced palatal fricative	n.a.	ç/ ̊, ç/ ̃, ç/¬	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a.	n.a.
ʝ	42	Fricative	Voiced	Palatal	Voiced palatal fricative	n.a.	ʝ/ ̊, ʝ/ ̃, ʝ/¬	p, b, t, d, f, v, ʊ s, n, l	k, g	n.a.	n.a.
x	43	Fricative	Unvoiced	Velar	Unvoiced velar fricative	n.a.	x/ ̊, x/ ̃, x/¬	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a.	n.a.
ɣ	44	Fricative	Voiced	Velar	Voiced velar fricative	n.a.	ɣ/ ̊, ɣ/ ̃, ɣ/¬	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a.	n.a.
χ	45	Fricative	Unvoiced	Uvular	Unvoiced uvular fricative	n.a.	χ/ ̊, χ/ ̃, χ /¬	Always oral consonant error	Always oral consonant error	n.a.	n.a.
ʁ	46	Fricative	Voiced	Uvular	Voiced uvular fricative	n.a.	ʁ/ ̊, ʁ/ ̃, ʁ/¬	Always oral consonant error	Always oral consonant error	n.a.	n.a.

Sound	Idx	Manner of articulation	Voicing	Place of articulation	Text to appear in SPEAK	Correct production if:	VPI symptoms present if:	Oral consonant error if:	Not oral consonant error if:	Non-oral consonant error if:	Not non-oral consonant error if:
ħ	47	Fricative	Unvoiced	Pharyngeal	Unvoiced pharyngeal fricative	n.a.	n.a.	n.a.	n.a.	always non-oral consonant error	always non-oral consonant error
ʕ	48	Fricative	Voiced	Pharyngeal	Voiced pharyngeal fricative	n.a.	n.a.	n.a.	n.a.	always non-oral consonant error	always non-oral consonant error
h	49	Fricative	Unvoiced	Glottal	Unvoiced glottal fricative	n.a.	n.a.	n.a.	n.a.	always non-oral consonant error	always non-oral consonant error
ɦ	50	Fricative	Voiced	Glottal	Voiced glottal fricative	n.a.	n.a.	n.a.	n.a.	always non-oral consonant error	always non-oral consonant error
ɬ	51	Lateral, fricative	Unvoiced	Alveolar	Unvoiced alveolar lateral fricative	n.a.	ɬ/ ̊, ɬ/˜, ɬ/¬	n.a.	n.a.	n.a.	n.a.
ɮ	52	Lateral, fricative	Voiced	Alveolar	Voiced alveolar lateral fricative	n.a.	ɮ/ ̊, ɮ/˜, ɮ/¬	n.a.	n.a.	n.a.	n.a.
ʋ	53	Approximant	Voiced	Labiodental	Voiced labiodental approximant	n.a.	ʋ̃, ʋ/ ̊, ʋ/˜	n.a.	n.a.	n.a.	n.a.
ɹ	54	Approximant	Voiced	Alveolar	Voiced alveolar approximant	n.a.	ɹ/ ̊, ɹ/˜	n.a.	n.a.	n.a.	n.a.
ɻ	55	Approximant	Voiced	Retroflex	Voiced retroflex approximant	n.a.	ɻ/ ̊, ɻ/˜	n.a.	n.a.	n.a.	n.a.
j	56	Approximant	Voiced	Palatal	Voiced palatal approximant	n.a.	j/ ̊, j/˜	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.
ɰ	57	Approximant	Voiced	Velar	Voiced velar approximant	n.a.	ɰ/ ̊, ɰ/˜	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.
l	58	Lateral approximant	Voiced	Alveolar	Voiced alveolar lateral approximant	n.a.	l/ ̊, l/˜	n.a.	n.a.	n.a.	n.a.
ɭ	59	Lateral approximant	Voiced	Retroflex	Voiced retroflex lateral approximant	n.a.	ɭ/ ̊, ɭ/˜	n.a.	n.a.	n.a.	n.a.

Sound	Idx	Manner of articulation	Voicing	Place of articulation	Text to appear in SPEAK	Correct production if:	VPI symptoms present if:	Oral consonant error if:	Not oral consonant error if:	Non-oral consonant error if:	Not non-oral consonant error if:
ʎ	60	Lateral approximant	Voiced	Palatal	Voiced palatal lateral approximant	n.a.	ʎ/ ̊, ʎ/ ̃	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.
ʟ	61	Lateral approximant	Voiced	Velar	Voiced velar lateral approximant	n.a.	ʟ/ ̊, ʟ/ ̃	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.
w	62	Approximant	Voiced	Velar labial	Voiced velar-labial approximant	n.a.	w/ ̊, w/ ̃	n.a.	n.a.	n.a.	n.a.
ɧ	63	Fricative	Unvoiced	Postalveolar and velar	Unvoiced postalveolar-velar fricative	n.a.	ɧ/ ̊, ɧ/ ̃, ɧ/¬	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.
Ø	64	Omission of sound	n.a.	n.a.	Sound omission	n.a.	n.a.	n.a.	n.a.	always non-oral consonant error	always non-oral consonant error
pʰ	65	Plosive	Unvoiced	Bilabial	Aspirated unvoiced bilabial plosive	n.a.	pʰ/ ̊, pʰ/ ̃, pʰ/¬	n.a.	n.a.	n.a.	n.a.
tʰ	66	Plosive	Unvoiced	Alveolar	Aspirated unvoiced alveolar plosive	n.a.	tʰ/ ̊, tʰ/ ̃, tʰ/¬	n.a.	n.a.	n.a.	n.a.
tˢ	68	Plosive	Unvoiced	Alveolar	Affricated unvoiced alveolar plosive	n.a.	tˢ/ ̊, tˢ/ ̃, tˢ/¬	n.a.	n.a.	n.a.	n.a.
kʰ	67	Plosive	Unvoiced	Velar	Aspirated unvoiced velar plosive	n.a.	kʰ/ ̊, kʰ/ ̃, kʰ/¬	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.
b̥	69	Plosive	Unvoiced	Bilabial	Nasal emission unvoiced bilabial plosive	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
d̥	70	Plosive	Unvoiced	Alveolar	Nasal emission unvoiced alveolar plosive	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
g̊	71	Plosive	Unvoiced	Velar	Nasal emission unvoiced velar plosive	n.a.	Always VPI-symptom	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.

Sound	Idx	Manner of articulation	Voicing	Place of articulation	Text to appear in SPEAK	Correct production if:	VPI symptoms present if:	Oral consonant error if:	Not oral consonant error if:	Non-oral consonant error if:	Not non-oral consonant error if:
p̊	72	Plosive	Unvoiced	Bilabial	Nasal emission unvoiced bilabial plosive	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
t̊	73	Plosive	Unvoiced	Alveolar	Nasal emission unvoiced alveolar plosive	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
k̊	74	Plosive	Unvoiced	Velar	Nasal emission unvoiced velar plosive	n.a.	Always VPI-symptom	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.
f̊	75	Fricative	Unvoiced	Labiodental	Nasal emission unvoiced labiodental fricative	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
s̊	76	Fricative	Unvoiced	Alveolar	Nasal emission unvoiced alveolar fricative	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
b̃	85	Plosive	Voiced	Bilabial	Nasalised voiced bilabial plosive	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
d̃	86	Plosive	Voiced	Alveolar	Nasalised voiced alveolar plosive	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
g̃	87	Plosive	Voiced	Velar	Nasalised voiced velar plosive	n.a.	Always VPI-symptom	p, b, t, d, s, f, v, ʊ, n, l	k, g	n.a.	n.a.
ʊ̃	88	Approximant	Voiced	Labiodental	Nasalised voiced labiodental approximant	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
n̪[(2)]	77[(2)]	Nasal	Voiced	Interdental	Interdental nasal	n.a.	d	n.a.	n.a.	p, b, t, g, k, f, s, v, ʊ, l	n.a.
t̪	79	Plosive	Unvoiced	Interdental	Interdental unvoiced plosive	n.a.	t̪/ʱ, t̪/˜, t̪/¬	n.a.	n.a.	n.a.	n.a.
d̪	78	Plosive	Voiced	Interdental	Interdental voiced plosive	n.a.	d̪/ʱ, d̪/˜, d̪/¬	n.a.	n.a.	n.a.	n.a.
θ̪	80	Fricative	Unvoiced	Interdental	Interdental unvoiced fricative	n.a.	θ̪ʱ, θ̪˜, θ̪¬	n.a.	n.a.	n.a.	n.a.

Sound	Idx	Manner of articulation	Voicing	Place of articulation	Text to appear in SPEAK	Correct production if:	VPI symptoms present if:	Oral consonant error if:	Not oral consonant error if:	Non-oral consonant error if:	Not non-oral consonant error if:
m̥̊	81	Active nasal fricative	Unvoiced	Bilabial	Bilabial active nasal fricative	n.a.	n.a.	n.a	n.a	always non-oral consonant error	always non-oral consonant error
n̥̊	82	Active nasal fricative	Unvoiced	Alveolar	Alveolar active nasal fricative	n.a.	n.a.	n.a	n.a.	always non-oral consonant error	always non-oral consonant error
ŋ̊̊	83	Active nasal fricative	Unvoiced	Velar	Velar active nasal fricative	n.a.	n.a.	n.a	n.a.	always non-oral consonant error	always non-oral consonant error
N̊̊	84	Active nasal fricative	Unvoiced	Unknown	Active nasal fricative, place unspecified	n.a.	n.a.	n.a	n.a	always non-oral consonant error	always non-oral consonant error
ʰ	89	Aspiration	n.a.	n.a.	Aspiration	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
̃	92	Nasalised	n.a.	n.a.	Nasalisation	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
̊̊	90	Nasal emission	n.a.	n.a.	Nasal emission	n.a.	Always VPI-symptom	n.a.	n.a.	n.a.	n.a.
̃̃	91	Denasalised	n.a.	n.a.	Denasalisation	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
¬	93	n.a.	n.a.	n.a.	Weak intra-oral pressure	n.a.	Always VPI symptom	n.a.	n.a.	n.a.	n.a.
b̥	94	Plosive	voiced	Bilabial	Unvoiced bilabial plosive	n.a.	b̥, b̃, b/ ̊̊, b/ ̃, b/¬	n.a.	n.a.	n.a.	n.a.
d̥	95	Plosive	voiced	Dental/ Alveolar	Unvoiced dental/alveolar plosive	n.a	d̥, d̃, d/ ̊̊, d/ ̃, d/¬	n.a.	n.a.	n.a.	n.a.
g̊̊	96	Plosive	voiced	Velar	Unvoiced velar plosive	n.a	g̊̊, g̃, g/ ̊̊, g/ ̃, g/¬	p, b, t, d, f, v, ʊ, s, n, l	k, g	n.a	n.a

Notes:
(1) Any sound in combination with a glottal (Idx 13) will be treated as a glottal only (Idx 13).
(2) Exceptions to the VPI symptoms interpretation (IDx: 14,15,16,17,19,77) which treats the Sounds and VPI symptoms present column the opposite way around, see *Section 18.5.2.1.2.*