Critical Review:

What is the relationship between nasalance as measured by the Nasometer and perceptual measures of nasality in English speaking children with a history of cleft palate and/or related craniofacial anomalies?

Elizabeth Hamilton
M.Cl.Sc (SLP) Candidate
Western University: School of Communication Sciences and Disorders

This critical review examines the evidence regarding the relationship between nasalance and nasality in English speaking children with a history of cleft palate and/or related craniofacial anomalies. A literature search using computerized databases yielded four studies that met the search criteria: three single group correlational studies, and one two group correlational study. Overall, the reviewed literature indicates a suggestive level of evidence that a substantially positive relationship exists between these two measures when non-nasal stimuli are used, and equivocal evidence of a significant relationship when mixed stimuli are used. Clinical implications are discussed.

Introduction

Children with cleft palate often present with hypernasal speech, most commonly due to a structural deficiency of the velum. Some reports indicate that 20 to 43% of these children will still have velopharyngeal insufficiency following surgical palatal repair, and may continue to sound hypernasal as a result (Kummer, Clark, Redle, Thomsen, & Billmire, 2012, p.146).

It is often the responsibility of the speech-language pathologist to determine the presence and severity of hypernasal speech in children with a history of cleft lip and palate or other related craniofacial anomalies. In some cases, the paediatric otolaryngologist may rely on the speech-language pathologist’s report when making recommendations regarding the surgical management of structural defects affecting a child’s speech.

Although current literature promotes perceptual evaluation by the speech-language pathologist as the “gold standard” for assessing nasal resonance (Sweeney & Sell, 2008, p.266), a recent survey of cleft palate/craniofacial professionals revealed that there appears to be no standard protocol for carrying out these perceptual assessments. Additionally, perceptual rating scales are inherently flawed due to factors such as variations in listener bias and experience (Kummer et al., 2012, p.150). As such, the American Cleft Palate-Craniofacial Association (ACPA) indicates the importance of including both perceptual and instrumental measures when conducting preoperative and postoperative speech assessments (ACPA, 2000). Still, recent studies suggest that although the majority of cleft lip and palate centers carry out perceptual evaluations, they are not consistently accompanied by instrumental measures.

The Nasometer (Kay Elemetrics 1986) is often used by speech-language pathologists to supplement the perceptual evaluation of nasal resonance. The Nasometer is a microcomputer-based system designed to measure oral and nasal acoustic sound signals through calculating a score, which represents the ratio of the energy in the two signals (Sweeney & Sell, 2008, p.267). It is a non-invasive, indirect method of obtaining objective data regarding resonance.

Many researchers have investigated the relationship between perceptual ratings of nasality and nasalance as measured by the Nasometer by conducting correlational studies. In these studies, mean nasalance scores for each participant’s oral recitations are captured by the Nasometer. Perceptual ratings are then assigned by one or more listeners. Finally, nasalance and perceptual values are analyzed, and a strength of relationship is determined by using a statistical measure of association. Perhaps the most familiar measure of calculating this relationship is by using the Pearson product-moment correlation coefficient, which measures the strength of linear dependence between two variables. Nasometer test sensitivity and specificity are additional statistical measures often included in these studies. Test sensitivity relates to the test’s ability to identify positive results, (in this case, hypernasality) and test specificity relates to the test’s ability to identify negative results, (participants who are not hypernasal).

Despite widespread use of the Nasometer in clinical and research settings, studies that have examined the relationship between perceptual measures of nasality and nasalance as measured by the Nasometer have yielded conflicting results (Sweeney & Sell, 2008, p.267).
Objectives

The primary objective of this review is to critically evaluate the existing literature that examines the relationship between nasalance as measured by the Nasometer and perceptual measures of nasality in English speaking children with a history of cleft palate and/or related craniofacial anomalies. The secondary objective of this review is to provide evidence-based recommendations for the comprehensive clinical assessment of resonance in the above population.

Methods

Search Strategy
Computerized databases including PubMed, JSTOR, and Cochrane Library were searched. The following key terms were targeted: (nasalance OR Nasometer) AND (nasality OR perceptual) AND (cleft palate). The search was limited to articles written in English. An examination of articles cited within the retrieved articles revealed additional studies for review.

Selection Criteria
Studies selected for review were required to examine the relationship between the perceptual evaluation of nasality and nasalance as measured by the Nasometer in the child population only. Only studies that included subjects with a history of cleft palate and/or related craniofacial anomalies were included for review. No limits were placed on the method of perceptual ratings of nasality.

Data Collection
The literature search yielded four articles that met the selection criteria described above. These articles consisted of three single group correlational studies, and one two group correlational study.

Results

Watterson, McFarlane, and Wright (1993) conducted a single group, correlational study with the primary purpose of evaluating the relationship between nasalance and hypernasality. The secondary purpose of their study was to examine the possible influence of nasal consonants on the above relationship. The subjects included in this study were 25 children between the ages of 3;4 and 13;0. All children were followed by the Northern Nevada Craniofacial Team and presented with a broad range of nasality.

Watterson et al. (1993) determined that stimulus passages provided by Fletcher et al. (1989), (The Rainbow Passage, The Zoo Passage, and a series of nasal sentences) were generally too difficult for children to recite fluently. Instead, the authors created simplified passages that contained nasal phonemes in a proportion similar to the original passages of Fletcher et al: the “non-nasal” passage contained no nasal phonemes, the “nasal” passage contained approximately 35% nasal phonemes, and the “standard” passage contained a proportion of nasal phonemes similar to their combined frequency of occurrence in Standard English, (10%). All participants were required to repeat the passages phrase by phrase, following the examiner model.

The ten listener judges included six practicing speech-language pathologists and four speech-language pathology graduate students. All were experienced in the evaluation of speech disorders associated with cleft palate. The judges were required to rate each speech sample for nasality using a five-point scale. A rating of “one” represented normal nasalance, and a rating of “five” represented severe hypernasality.

The 25 median nasality ratings associated with each of the three passages were correlated with the 25 mean nasalance scores. For the non-nasal passage, the obtained correlation coefficient between nasality and nasalance was 0.73, but identified a number of the hypernasal subjects as normal (sensitivity = 0.42). The nasal passage was identified as deviant by the Nasometer, (specificity = 0.55). The standard passage had a good agreement with the listeners in identifying normal resonance (specificity = 0.73), but identified a number of the hypernasal subjects as normal (sensitivity = 0.42). The nasal passage did not identify any of the subjects as hypernasal. The authors concluded that the Nasometer may provide meaningful data for some patients but not for others, and urged clinicians to use their judgments with regards to whether nasometric measures are useful or not (Watterson et al., 1993, p.25).

In this article, Watterson et al. (1993) included a level of detail that would provide other researchers with enough information to replicate their procedures as closely as possible, including the operation and positioning of components of the Nasometer. Additional strengths of the article include a detailed and appropriate statistical analysis of the data, as well as a comprehensive
discussion of variables that may have influenced the results and led to weak correlational values. Although participants were screened for sensorineural hearing loss, of potential concern is the fact that the authors did not screen participants for syndromes associated with craniofacial anomalies. Several recent studies indicate that individuals with cleft palate in addition to a diagnosis of a syndrome often have several anatomic and physiologic differences that affect velopharyngeal closure (Widdershoven, Stubenitsky, Breugem, & MinkvanderMolen, 2008). Consequently, results of the Watterson et al. (1993) study may not adequately reflect the vocal resonance of children with isolated cleft palate, or the resonance of children with syndromes that do not affect velopharyngeal closure. It is also important to note that perceptual evaluations of nasality were made based on an audio recording of passage recitations. It is possible that this reduced the clinical validity of the correlational results, as live ratings are often used in clinical protocols (Sell, 2004).

Watterson, Hinton, and McFarlane (1996) conducted a two group, correlational study primarily to evaluate the use of novel stimuli in the assessment of nasalance. They also examined the correlation between listener judgments of hypernasality and nasalance scores, which will remain the focus of this discussion. The subjects were 20 children ranging in age from 3:0 to 6:6 who were followed on a regular basis by the Northern Nevada Craniofacial Team and presented with a broad range of nasality.

The authors of this study designed two novel stimuli passages that they deemed syntactically and semantically appropriate for young children: the Turtle Passage, which contained no nasal phonemes, and the Mouse Passage, which contained approximately 11% nasal phonemes. The Zoo Passage (Fletcher et al., 1989) was also included as a stimulus. All participants were asked to recite the passages phrase by phrase, following a model provided by the examiner.

The eight listeners assigned to provide perceptual evaluations were all speech-language pathology graduate students who had training in the assessment and treatment of speech disorders relating to cleft palate. The judges were required to rate each speech sample for nasality using a five-point scale. A rating of “one” represented normal nasal resonance, and a rating of “five” represented severe hypernasality.

Mean nasality ratings and mean nasalance scores were analyzed using the Pearson correlation coefficient. A correlation coefficient of $r = 0.70$ ($p < .0001$) was obtained for the Zoo Passage, described by the authors as a ‘substantial relationship’ between nasalance and nasality. For the Turtle Passage, a correlation coefficient of $r = 0.51$ ($p < .05$) was obtained, also considered a substantial relationship. The Pearson correlation coefficient for the Mouse passage was $r = 0.32$ ($p = .17$) which was not considered significant. Nasometer test sensitivity and specificity measures for the Zoo Passage revealed a sensitivity score of 0.72 and a specificity score of 0.50. For the Turtle passage, a sensitivity score of 0.83 was obtained, with a specificity score of 0.0. An in-depth analysis of sensitivity and specificity scores by the authors revealed an ambiguity in classifying children with borderline hypernasality, both in terms of perceived nasality and nasalance as measured by the Nasometer. Watterson et al. (1996) concluded that clinicians can be most confident in nasalance scores for individuals who are “obviously normal or obviously hypernasal”, and least confident in nasalance scores for patients who are borderline-normal.

By using their own population to establish normal nasalance values for their stimuli, and by separately analyzing the results of male and female speakers, Watterson et al. (1996) were able to control for variables that have historically been questioned to influence results of nasalance and nasality measures. For instance, after reporting regional differences in nasalance, Seaver, Dalston, Leeper, and Adams (1991) suggested that it might be necessary to establish “regional norms”. Fletcher (1978) found differing nasalance values between school-aged boys and girls; however, in a later study reported no sex differences (1989). Additional strengths of this study include the careful construction of novel stimuli by the authors, as well as their decision to exclude subjects with articulation errors, thereby reducing potential inflation of nasalance scores. Unfortunately, the authors did not screen participants for hearing loss or craniofacial syndromes, both of which are factors known to influence nasal resonance (Widdershoven et al, 2008).

Sweeney and Sell (2008) conducted a single group, correlational study that compared a clinician’s perceptual ratings with Nasometry measurements. The group of children who participated in the study consisted of a consecutive series of 50 children who had been referred to a national cleft lip and palate clinic for investigation of speech problems. There were 30 males and 20 females, ranging in age from 4:10 to 15:10. The participants presented with normal nasality, a range of hypernasality, or hyponasality. For the purposes of this review, only the data collected from participants who were hypernasal will be considered.

The speech samples used for the perceptual evaluation included: single words and syllables, adapted sentences from the GOS.SP.ASS (Sell, Harding, & Grunwell,
1999); automatic speech (‘Jack and Jill’, counting from one to 20 and from 60 to 70) and a minimum of two minutes of conversational speech. The GOS.SP.ASS sentences were adapted to include low-pressure consonants, and sentences were reordered according to consonant type. Further adaptation was made to the GOS.SP.ASS sentences to ensure that the total test sentences contained a proportion of nasal consonants similar to the distribution of nasal consonants in English conversational speech.

Live perceptual ratings were provided by a specialist speech and language therapist, using the Temple Street Scale of Nasality and Nasal Airflow Errors (Temple Street Scale), developed by Triona Sweeney. On this 6-point scale, a score of “zero” indicates absent hypernasality, and a score of “five” indicates severe hypernasality. In this study, each participant was asked to repeat the 16 test sentences following the examiner.

The Pearson product moment correlation was used to calculate the relationship between nasalance scores and perceptual ratings of nasality. All correlation analyses were two tailed (p = 0.05). Results of the correlation analyses indicated a ‘substantial positive relationship’ between perceptual ratings of hypernasality and nasalance scores on both the total test sentences, (r = 0.74, p < 0.001) and the high-pressure consonant sentences, (r = 0.74, p < 0.001). In terms of the low-pressure consonant sentences, a ‘weaker but significant’ relationship was found, (r = 0.69, p < 0.001). For the total test sentences, using a cut-off of 35%, sensitivity was 50.83, and specificity was 50.78. In order to calculate an overall efficiency rating, the authors added the number of times the perceptual ratings agreed with scores from the Nasometer, and then divided by the total number of participants tested. Using this method, an efficiency value of 50.82 was calculated, and overall it was stated that there was a ‘good relationship’ between perceptual ratings of the total test sentences and the nasometric values. Sensitivity, specificity, and efficiency calculations for the high-pressure consonant sentences, (cut-off = 24%, sensitivity = 50.83, specificity = 50.86, efficiency = 50.84) and the low-pressure consonant sentences (cut-off = 28%, sensitivity = 50.88, specificity = 50.78, efficiency = 50.86) also indicated a good relationship between perceptions of nasality and nasalance scores. Sweeney and Sell (2008) concluded that professionals can have confidence in the clinical findings when there is agreement between perceptual and nasometric measurements. They highlighted the importance of use of the Nasometer to supplement but never replace perceptual judgment.

Sweeney and Sell (2008)’s article included thorough background information and a sound rationale to support their research statement. Additional strengths of the study included systematic measures of reliability, the inclusion of all participant data on all tasks, appropriate statistical analyses, and a detailed discussion of the results. Interestingly, instead of having perceptual raters use EAI or DME measurements, the authors provided listeners with the Temple Street Scale of Nasality and Nasal Airflow Errors, which is a descriptive scale for the assessment of nasality. The clear definition of terms within this rating scale may have strengthened test-retest and inter-judge reliability (Wirz & MacKenzie Beck, 1995). Although inclusionary and exclusionary criteria were described in detail, the authors’ decision to include participants with syndromes and articulatory errors unfortunately limits the extent to which the results of the study can be compared across other similar studies.

Brancamp, Lewis, & Watterson (2010) conducted a single group, correlational study in order to assess the nasalance/nasality relationship and Nasometer test sensitivity and specificity when nasality ratings were obtained with both equal appearing interval (EAI) and direct magnitude estimation (DME) scaling procedures. The participants in this study were 39 children and adolescents between the ages of 3;8 to 17;2. 25 of these children had a history of hypernasality and were followed by a cleft lip and palate team, while the remaining 14 had no history of hypernasality or speech and language problems.

Participants were required to recite the Turtle Passage, (a passage devoid of nasal consonants) which was composed by Watterson et al. in 1996 for use as a child-friendly alternative to the Zoo Passage. Nasalance values were obtained from the Nasometer, and recitations were recorded for future perceptual evaluations. The perceptual evaluations were provided by a judge with more than 30 years of experience in assessing resonance disorders in the cleft palate population. The judge was required to complete both an EAI and a DME scaling procedure for each speech sample. For the DME procedure, ratings between one and ten were defined as normal, and ratings higher than ten were defined as hypernasal. Ratings were made in reference to a moderately hypernasal speech sample assigned a rating of 100. The EAI measure was a five-point scale in which a score of “one” represented normal resonance and a score of “five” represented severe hypernasality.

Separate bivariate correlations were calculated in order to assess the strength of relationship between nasalance scores and nasality ratings using EAI scaling as well as the relationship between nasalance and nasality ratings using DME scaling. A moderate correlation was found
between both nasalance scores and EAI nasality ratings, \( r = .63, p < .01 \) and nasalance scores and DME nasality ratings, \( r = .59, p < .01 \). These values represented a substantial relationship in both cases, and a difference test showed no significant difference between the coefficients, \( (t(36) = -.50, p > .05) \). For EAI nasality ratings, sensitivity was .71 and specificity was .73, with an overall efficiency value of .72. For DME nasality ratings, sensitivity was .62 and specificity was .70, with an overall efficiency value of .64. Brancamp et al. (2010) concluded that clinicians should be able to obtain reliable and valid nasality estimates using either EAI or DME scaling.

Despite the authors’ careful analyses of their data and useful results regarding EAI and DME scaling procedures, Brancamp et al.’s (2010) study does not provide the reader with sufficient information regarding the relationship between nasality ratings and nasalance measures of cleft palate speech. One obvious shortcoming of the article is that the authors did not publish the data or results of the 25 participants with a history of hypernasality separately from the 14 participants without resonance disorders. Additionally, the decision to use the perceptual ratings of only one individual may have affected the reliability of the results, as the single listener’s judgments may not be representative of the perceptual abilities of all speech-language pathologists.

**Discussion**

Overall, the examined research provides suggestive evidence of a substantial positive relationship between nasalance and nasality in children with cleft palate and/or related craniofacial anomalies when tested with non-nasal stimuli. Equivocal evidence of a significant positive relationship was provided when participants were tested with passages containing both oral and nasal consonants. Due to the variations in test procedures across the examined studies, the results should be interpreted with caution.

In the context of these studies, it is clear that choice of stimuli is one factor that can significantly impact nasalance and nasality measures, as well as the results of correlational measures. Nasalance scores from non-nasal passage recitations generally yielded the strongest relationships with perceptual ratings, with correlation values ranging from 0.49 \( (p = .006) \) in the Watterson et al. (1993) study, to .70 \( (p < 0.0001) \) in the Watterson et al. (1996) study. Only two of the four articles reviewed included passages that contained both nasal and oral consonants in their stimuli, (Watterson et al. 1993; Sweeney & Sell, 2008). While results of the Sweeney and Sell (2008) article indicated substantial positive relationships between perceptual judgments and nasalance scores for mixed sentences \( (0.74, p < .001 \) for both high-pressure and low-pressure sentences), Watterson et al.’s (1993) results indicated a weak relationship, \( (0.24, p = .13) \). Watterson et al.’s (1993; 1996) conclusion that nasal sentences are not useful in identifying children who are hypernasal is consistent with other studies that have examined hypernasality in various other populations. It is perhaps due to these findings that the authors of the remaining papers in this review excluded nasally loaded passages altogether.

When attempting to compare results across studies, one must also consider the criteria that were used for participant selection. Individuals with cleft lip and palate represent a highly heterogeneous group; for instance, the severity of an individual’s palatal cleft can range from an isolated bifid uvula with perceptually normal resonance, to a full bilateral cleft of the lip and palate, resulting in severe hypernasality accompanied by nasal air emissions. Although none of the reviewed studies controlled for nasal air emissions, certain literature suggests that mean nasalance scores may be affected (Karnell, 1995; Karnell, 2011). Individuals with hearing loss and/or known syndromes may demonstrate increased hypernasality when compared to individuals with isolated clefts (Widdershoven et al., 2008). Additionally, some researchers have argued that those with articulation errors should be excluded from studies such as these due to potential inflation of nasalance scores. The above inclusionary criteria were not consistent across the four studies, limiting the ability to compare across studies and the overall reliability of results.

Finally, an additional area of controversy when examining the nasalance/nasality relationship is the type of rating scale used during perceptual evaluations. Equal appearing interval scales and rating scales that use descriptive category judgments, (e.g., mild, moderate, and severe) are often used; however, the number of points in these scales has ranged from four to nine, or even 11 points (Whitehill, 2004). A rating scale with a small range may lead to important perceptual information being overlooked, while a scale with a large range may take away from the rater’s ability to focus on what they are hearing and select an accurate rating from multiple options. All of the studies considered for this review used a five-point scale except for the Sweeney and Sell (2008) article, which used a six-point scale. This scale also included descriptors, (e.g., a “two” represented mild/moderate hypernasality, characterized as ‘unacceptable distortion, evident on high vowels’) which perhaps increased reliability, but creates difficulty when attempting to compare results across studies. Overall, the lack of definition of terms used in
the scales across studies may have resulted in poor test retest and inter-judge reliability (Sweeney & Sell, 2008).

**Conclusion and Recommendations**

Collectively, the studies reviewed offer a suggestive level of evidence supporting a substantial positive relationship between nasalance as measured by the Nasometer and perceptual measures of nasality in English speaking children with a history of cleft palate and/or related craniofacial anomalies when non-nasal passages are used as stimuli. The two studies that included both nasal and oral consonant passages in their stimuli offer equivocal evidence of a significant relationship. Factors such as choice of stimuli, inclusionary criteria, and type of perceptual rating scale used the studies may have influenced results, also limiting the ability to compare across studies. While more research is needed in order to determine appropriate stimuli as well as to develop a suitable perceptual rating scale, there is also a need for consistency across test procedures. Such consistency would allow for a consensus to be made regarding the true nature and extent of the relationship between nasalance and nasality in the cleft palate population.

**Clinical Implications**

Due to the variable evidence provided by the articles included in this review as well as the homogeneous nature of the cleft palate population, clinicians are encouraged to integrate their own clinical expertise with the latest available external evidence when considering the validity and reliability of perceptual and nasometric results. Although it is recommended that a comprehensive evaluation of resonance consist of both perceptual and instrumental measures, clinicians should never rely solely on results of the Nasometer when making decisions regarding management or treatment.

**References**


