The Nasal Width and Boxiness Index: Introduction and Pilot Study on Reliability and Validity of Sonographic Morphometry

Abel-Jan Tasman, MD1 Ralph Litschel, MD2

1 Department of ENT, Cantonal Hospital St. Gallen, St. Gallen, Switzerland
2 Private Practice, St. Gallen, Switzerland


Abstract

Both intended and unintended surgical modifications of nasal width and shape of the nasal tip continue to be of interest to the rhinoplasty surgeon. As validated instruments for quantifying width and boxiness are lacking, the objective of this study was to introduce a width index and a boxiness index for the nasal dorsum and the nasal tip. A width index and a boxiness index were defined within the methodological limits of noncontact sonography. The reliability of both indices was studied by comparing the measurements of two examiners on the noses of five volunteers. The validity of the indices was studied by correlating the sonographic width and boxiness with the 5-point Likert scale ratings of photographs of 5 noses by 21 lay persons. Nasal width was defined as the diameter at a distance of 5 mm from the skin surface on a sonographic cross-section perpendicular to the skin surface. Boxiness was defined as the quotient of width at a depth of 1 and 5 mm. Bland–Altman analysis revealed negligible bias between both examiners and 95% of limits of agreement of 13, 7, and 13% for width at 1 mm, width at 5 mm, and boxiness, respectively. Corresponding Pearson’s correlation coefficients were $r = 0.93$, $r = 0.93$, and $r = 0.71$. The correlation between the cumulative lay persons’ scores and sonographic width and boxiness were $r = 0.97$, $r = 0.66$, and $r = 0.81$ for nasal tip width, dorsal width, and boxiness, respectively. Both the width at a depth of 5 mm as measured with sonography and the boxiness index that is defined as width at a depth of 1 mm divided by the width at a depth of 5 mm may prove to be acceptable surrogate parameters for width and boxiness of the nose in comparative morphometric studies.

Keywords
► rhinoplasty
► morphometry
► sonography
► nasal width

The literature holds a plethora of rhinoplasty techniques, but articles based on morphometric studies are sparse. Decades ago, techniques for studying facial morphology that have been applied widely in the analysis of facial dysmorphology used direct anthropometry with a ruler, calipers, tape measure, and protractor.1 For nasal anatomy, several technologies have been used for anthropometric measurements: a sliding caliper and the estimation of anthropometric points.2,3 Linear indices have been calculated for the nose such as the index of prominence to width, defined as the interalar distance.4 A source of error in these studies was observer error in locating the landmarks with insufficient reliability of identification of soft tissue landmarks on photogrammetry.5 Two studies that aimed at morphometric assessment measured the effect of spreader grafts on nasal dorsal width on two-dimensional (2D) photographs of patients based on estimating dorsal width on the photographs.6,7 In these studies, a line representing the perceived width was drawn or marked points on perceived dorsal aesthetic lines were connected, respectively. The authors of
both studies did not provide information on the reliability of the measurements or on the specific criteria used to define anthropometric points. This is illustrated by the exclusion of 17% of the patients in one study due to differences in light reflections and in skin brightness before and after surgery.7

In modern anthropometry of such complex structures as the face, classic direct anthropometry is being coupled to or replaced by three-dimensional (3D) image analyzers (laser scanners, 3D range-cameras, optoelectronic instruments, stereophotogrammetry, Moiré topography) as well as contact instruments (electromagnetic and electromechanical digitizers, ultrasound probes).8–13 Three-dimensional scanning techniques in particular have become more and more common.14 Before any system can be implemented in quantitative studies of patient populations, the error in producing an image and error in taking measures from the images produced must be evaluated.15 Precision of measurements using the anatomical points obtained by 3D scanners has been described as more than sufficient for clinical needs and greater than that of other methods, such as direct anthropometry and 2D photography. In addition, 3D models have not only provided a high level of technical precision but also a sufficient intra- and interobserver reliability regarding landmark identification.16,17 Estimates of intermethod reliability as defined by the correlation coefficient comparing caliper to 3D image-based measurements for all participants were 0.4 to 0.75 for measurements of relatively small dimensions, for example, columellar width and mean absolute differences of less than 1 mm, relative error measurement less than 5 mm, and technical error of measurement less than 1 mm.18,19 However, correlation measures the strength of a relation between two variables, not the agreement between them. Bland and Altman therefore made the point that the correlation coefficient does not help in interpreting a clinical measurement on a given patient.20 For this purpose, the repeatability coefficient should be calculated instead, which is the difference that will be exceeded by only 5% of pairs of measurements on the same subject.21 Repeated measurements by the same method or measurements by different observers can be analyzed in a manner analogous to the limits of agreement. The main difference being that the average difference should be zero.22

Using 3D surface imaging, the preprocedural and acute postprocedural nasofrontal, nasofacial, nasolabial, and namental angles, early changes in the height and length of the nose, and nasal volume after filler injections have been compared.23 Also, 3D stereophotogrammetry (Vectra) and volumetric analysis (Geomagic) have shown that approximately two-thirds of edema after rhinoplasty resolves within the first month, 95% resolves after 6 months, and 97.5% resolves after 1 year.24 Off-the-shelf stereo cameras for the 3D assessment of morphometric variations caused by rhinoplasty were found to be useful for stereo-photogrammetric facial digitization and morphometric analysis of the human face.25 Regarding the spatial resolution of a 3dMD scanning system, a global error with a range from 0.1 to 0.5 mm was found in an experimental setting.14 Differences between direct measurements and 3D measurements of up to 1 mm were found when measuring nasal width.14,15,26 Some of the classical landmarks that examiners of real patients would tend to identify using palpation of underlying bony structures were found to be unreliable on 3D photogrammetry with error greater than 1 mm for the nasion and 2 mm for the glabella.15,27 Ultrasonography with a high-frequency transducer, by contrast, offers a spatial resolution of less than 0.1 mm and a cross-sectional image of soft tissues that has been used for morphometric assessment of the nasal surface and soft tissues as an adjunct to photogrammetric analysis and palpation for both surface and cross-sectional anatomy of the nose and for morphometry of dorsal cartilage grafts.9–14 More recently, sonography was used to study the correlation between the classification of skin thickness at the nasolabial fold as determined by the skinpinch test and sonographic measurements of the nasal skin.28,29 Clinical correlations and standard values by measuring 13 length and 4 ratios have been defined for nasal soft tissue and cartilaginous structures by means of high-resolution ultrasonography.14 Given the need for quantitative morphometric assessment of surgical techniques, this study aimed at defining sonographic criteria for nasal width and boxiness. Additional objectives were a preliminary study of the reliability and validity of sonographic measurement of the nasal tip and nasal dorsum width.

Methods

Sonographic Morphometry

A linear 33.7 mm, 7 to 18 MHz transducer (SP10–16-D, GE Healthcare) was used with a GE Voluson E8 sonograph (GE Healthcare). A noncontact method was used as described elsewhere9 (Fig. 1). The transducer was held at a 90° angle to the sagittal plane and a 90° angle to the skin surface. Measurements were taken in the midline at the tip defining

![Fig. 1](image-url) Noncontact sonography of the nasal tip. Ultrasound coupling gel bridges the surface of the sound probe and the nasal skin.
point, at the interpupillary line, and at a third point in the middle between the two previous points (Fig. 2). All measurements were made by the examiner at the time of sonography on still images using the caliper instrument with a resolution of 0.1 mm. Measurements were then transferred to a data file for analysis.

Reliability
Two examiners (A.J.T. and R.L.) both studied 5 volunteers. Three measurements were made, first at the rhinion, next at a point halfway between the rhinion and the tip, and finally at the tip in a plane containing the tip-defining points. Still images were stored and the width was measured at a depth of 1 and 5 mm for each measuring point. A second set of measurements was then taken in the same way by both examiners. Differences between the first and the second measurements were calculated. The measurements of the two examiners were compared by calculating the average of the two corresponding measurements of each examiner. These calculated means were then entered into a Bland and Altman plot analysis. In addition, Pearson's correlation was calculated.

Validity
The validity of both the width index and the boxiness index was studied by comparing lay persons' ratings of width and boxiness with sonographic data of 5 noses that clinically ranged from narrow to wide and from pointed to boxy. Twenty-one lay persons rated the width of the dorsum and the tip as well as boxiness of the tip (Figs. 2 and 3). A 5-point Likert scale ranging from 1 for very narrow or very pointed, to 5 for very wide or very boxy was used. Lay persons' average impression of width and boxiness was defined as the sum of all ratings for the dorsum and the tip and for tip boxiness. The correlation between the lay persons' cumulative scores and the sonographic indices was calculated.

Approval of the local ethics committee was obtained.

Results
Width and Boxiness Index
The width index of the nose was defined as the distance between opposing skin surface points measured on a sonographic cross-section of the nose, parallel to and at 5.0 mm

![Fig. 2 Location of sonographic measurements at the tip, the middle third, and the upper third. These figures and the Likert scale were used to instruct the lay raters on which anatomical location to judge regarding width and how to rate tip boxiness. The red, green, and blue circles on the photographs guided the lay raters regarding the anatomical location—upper and middle third and nasal tip respectively—to be evaluated and rated on the Likert scale on the right. The contour of the nasal tip in yellow on the base view of the nose referred to the schematic contours on the right, ranging from very boxy to very pointed.](image)

![Fig. 3 Base view of the five noses that were rated by lay persons regarding boxiness.](image)
The boxiness index of the nose was defined as the width measured at 1 mm from the surface, divided by the width measured at 5 mm from the surface (Fig. 4). The theoretical range of the calculated boxiness therefore was between 0.2 and 1.0 for a geometric triangular and rectangular cross-sectional shape.

All measurements could be taken in all subjects by both examiners. No data were discarded.

**Reliability**

Bland–Altman analysis of data for width at 1 mm calculated an average bias of 0.4% with 95% limits of agreement between −12.6 and 13.4% (Fig. 5). Pearson’s correlation between the first and second measurement was \( r = 0.92 \) and \( r = 0.90 \) for examiner 1 and examiner 2, respectively. The correlation between the average of examiner 1 and examiner 2 was \( r = 0.93 \).

Bland–Altman analysis of data for boxiness calculated an average bias of −1.2% with 95% limits of agreement between −8.2 and 5.8% (Fig. 6). Pearson’s correlation between the first and second measurement was \( r = 0.97 \) and \( r = 0.98 \) for examiner 1 and examiner 2, respectively. The correlation between the average of examiner 1 and examiner 2 was \( r = 0.93 \).
Bland–Altman analysis of data for boxiness calculated an average bias of 1.6% with 95% limits of agreement between −11.4 and 14.6% (Fig. 7). Pearson’s correlation between the first and second calculated boxiness index was \( r = 0.70 \) and \( r = 0.53 \) for examiner 1 and examiner 2, respectively. The correlation between the average of examiner 1 and examiner 2 was \( r = 0.71 \).

**Validity**
The correlation between sonomorphometric indices and cumulative lay persons’ ratings for the width of the nasal tip and for the middle third was \( r = 0.97 \) and \( r = 0.66 \), respectively (Figs. 8 and 9). The corresponding correlation for tip boxiness was \( r = 0.81 \) (Fig. 10). Face validity is illustrated by a clinical example (Fig. 11).

**Discussion**
Maintenance of middle third width after hump removal and shaping of the nasal tip with suture techniques and grafts continue to be topics of clinical interest. Spreader grafts and variations of folded upper lateral cartilages, often referred to as autospreader grafts or spreader flaps, reverse spreader grafts, and a spring graft have been advocated to prevent an unwanted narrowing of the middle third of the nasal dorsum.30–33 Alternative techniques such as the endonasal approach of middle vault reconstruction for the creation of a smooth dorsal profile and nasal valve reconstruction have also been proposed.34,35 Putative methodological limitations in studying these effects are the subjective nature of measuring width on portrait photographs. Two reports on the outcome of measurements on standard photographs did not provide information on the reliability or validity of the data.6,7 The difficulty and subjectivity of measuring was illustrated by the deletion of photographic images that were flawed by motion artifacts or inappropriate head position and selection of images that were deemed acceptable.15 Measuring outcome, such as intended or unwanted alteration of tip or dorsum width, requires a measuring instrument that ideally is widely available, reliable, and valid. Three-dimensional photogrammetric measuring systems may prove to have a sufficient spatial resolution to provide

Fig. 7 Bland and Altman plot analysis of the tip boxiness index.

Fig. 8 Lay persons’ perception of tip width and sonographic tip width at 5 mm. The size of the circles reflects the sum of the ratings for “very narrow” (−2) to very wide (2) (see Fig. 2).
examiner-independent data that will allow a quantification of clinically relevant changes in nasal width and shape of the nasal tip. Published studies based on 3D photography of rhinoplasty outcomes as yet do not provide clear information on the reproducibility of measurements in the submillimeter range. Also, access to these systems may still be limited. Computed tomography imaging by contrast is widely available and may be used for morphometric studies with a sufficient spatial resolution. Cost and exposing patients to ionizing radiation in particular will be prohibitive in most clinical settings. Sonography is widely available in many academic institutions and has been used for morphometric studies of nasal anatomy and surgical outcome.\textsuperscript{9,11–13} Advantages are widespread availability, high patient acceptance, superior spatial resolution, and unique visualization of cartilaginous nasal structures.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{Fig_9.png}
\caption{Lay persons’ perception of tip width and sonographic width of the middle third at 5 mm. The size of the circles reflects the sum of the ratings for “very narrow” (see \textsuperscript{9} Fig. 2).}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{Fig_10.png}
\caption{Lay persons’ perception of tip boxiness. The size of the circles reflects the sum of the ratings for “very boxy” (<2) to very pointed (2), as graphically illustrated in \textsuperscript{9} Fig. 2.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{Fig_11.png}
\caption{Clinical example of increased dorsal width after a hump reduction. Width of the middle third at 5 mm measured 11.5 mm before surgery (left) and 14.3 mm after hump reduction (right).}
\end{figure}
Studying the reliability and validity of width at distances of more than 5 mm would have been desirable. These may or may not reflect the perception of width even better. The maximum distance appeared to be limited by the properties of the ultrasound gel, which was sufficiently viscous to bridge the distance between the ultrasound probe and the incongruent surface of the nasal skin, yet liquefied enough to keep the surface anatomy unaltered. The width index was defined based on both the intuition of the authors and on the methodological feasibility and is herein presented as one option to quantify width. Depending on the methodology, the width of the nose at different depth levels may be appropriate. To the knowledge of the authors, nasal width and boxiness have not been defined in a similar way before. Also, published studies measuring changes in width have not measured the reliability or the validity of the method that was used. To assess the degree of agreement of the two methods, the correct statistical approach is not obvious. Correlation coefficients indicate the agreement between two measures rather than the differences. A superior method to assess if two measuring methods offer the same values, or two examiners obtain identical readings with one examining the second method or the correct statistical approach is not obvious. Also, published studies measuring changes in width have not measured the reliability or the validity of the method that was used. To assess the degree of agreement of the two methods, the correct statistical approach is not obvious. Correlation coefficients indicate the agreement between two measures rather than the differences. A superior method to assess if two measuring methods offer the same values, or two examiners obtain identical readings with one examining the second method compared with the first or the second measurement compared with the first. The Bland and Altman plot analysis. This analysis is a simple way of calculating a bias between the mean differences. It estimates the agreement interval containing 95% of the second method compared with the first or the second measurement compared with the first. The Bland and Altman analysis is primarily used to compare two different measuring instruments, but it may also be used to compare two examiners using the same method. In the latter case, average discrepancy between the two examiners (bias) should be close to zero. The analysis used in this study showed an average discrepancy between the two examiners that may be negligible in clinical settings. The limits of agreement appeared to be sufficiently narrow for the width readings at 1 and 5 mm, with a superior precision of the measurements at 5 mm which is probably due to a larger angle between the skin surface and the measuring line. This is reflected in higher correlation coefficients. Measurements at 1 mm were less reliable and valid, which is not surprising, given the smaller angle of the skin surface at 1 mm which affects precision of defining the skin surface. Reliability of boxiness, which is defined as the quotient of the 1 and 5 mm reading, is by nature less reliable than both the 1 and 5 mm width. Despite this inherent imprecision of the depth at 1 mm measurements and the boxiness index, lay persons’ ratings of both width at 5 mm and of boxiness correlated rather well with sonographic readings. This finding, while being far from robust, hints at a sufficient face validity and construct validity.

Conclusion

Both the width at a depth of 5 mm as measured with sonography and the boxiness index that is defined as width at a depth of 1 mm divided by the width at a depth of 5 mm may prove to be acceptable surrogate parameters for width and boxiness of the nose.

References