The Value of Averageness in Aesthetic Rhinoplasty: Humans Like Average Noses

Floris V.W.J. van Zijl, MD; David I. Perrett, PhD; Peter J.F.M. Lohuis, MD, PhD; Carolina E. Touw, MD; Dengke Xiao, PhD; and Frank R. Datema, MD, PhD

Drs van Zijl and Datema are ENT surgeons and Ms Touw is a medical student, Department of Otorhinolaryngology and Head and Neck Surgery, Erasmus Medical Center, Rotterdam, the Netherlands. Prof Perrett is a Professor and Mr Xiao is a Postdoctoral Research Fellow, School of Psychology and Neuroscience, University of St Andrews, St Andrews, UK. Dr Lohuis is an ENT surgeon, Department of Otorhinolaryngology and Head and Neck Surgery, University Hospital Sveti Duh, Zagreb, Croatia.

Corresponding Author: Dr Floris V.W.J. van Zijl, Department of Otorhinolaryngology and Head and Neck Surgery, Erasmus Medical Center, 's-Gravendijkwal 230, 3015 CE Rotterdam, The Netherlands
E-mail: f.vanzijl@erasmusmc.nl

Disclosures: The authors declared no potential conflicts of interest with respect to the research, authorship, and publication of this article.

Funding: The authors received no financial support for the research, authorship, and publication of this article.

Abstract

Background: The aesthetic ideal of the nose eludes clear definition. Averageness may be an important determinant of ideal nasal shape: research has shown that averageness plays an important role in the human perception of facial attractiveness.

Objectives: To test whether an averaged nasal shape is attractive, and whether deviation away from average is associated with decreased attractiveness.

Methods: Photographical series of the face were obtained from 80 Caucasian female volunteers aged 25-40 years. A mathematically averaged composite image was created using the first 40 volunteers. Forty-one panel members were recruited to judge the attractiveness of the nose of each original image and the composite, using a 5-point Likert scale ranging from 1 (very ugly) to 5 (very pretty). Deviation of nasal shape from average was calculated using a principal components analysis of standardized nasal landmarks.

Results: Twenty-one respondents were male (51%). Mean (SD) age of the respondents was 35.3 years ($\pm$15.6). The rating of the composite was significantly higher than the distribution of ratings for the 80 original nose images (4.2 vs 2.8, t=31.24, P<.001). The rating of the original nose images correlated negatively with deviation from average shape ($r = -.40$, n = 80, P<.001).

Conclusions: In Caucasian females, a mathematically averaged nose is an attractive nose. Furthermore, the more an individual nose shape resembles average shape, the more attractive it is rated. Calculating deviation from average before and after rhinoplasty may aid in objectively measuring aesthetic rhinoplasty outcome.

A comprehensive understanding of facial aesthetics is of paramount importance to rhinoplasty surgeons. In pursuit of high patient satisfaction, every rhinoplasty surgeon deals with the dilemma of what constitutes the ideal nose. Traditionally, neoclassical canons that focus on optimal alignment, angularity and proportionality of the face serve as guidelines. However, large cross-cultural anthropometrical studies show that only a minor percentage of the population exhibits these aesthetic ideals, and attractive faces do not conform to them any more than do less attractive faces.[1-3]

Evolutionary psychologists have studied facial attractiveness extensively, linking the importance of attractiveness to human interactions and sexual selection. They identified possible cues to what humans find universally attractive in others, such as averageness, symmetry, youthfulness, and sexual dimorphism.[4] Of these four factors, averageness has been proposed as the most important factor in perceiving facial attractiveness.[5] Averageness in this context does not refer to typical mean in the sense of 'common' or 'frequently occurring in the population', but more to a mathematically averaged, computer-manipulated composite of whole faces, that is based on mean numerical values across the face. In 1990, Langlois et. al produced such an 'average composite', and found that this composite was perceived as more attractive than the individual gray scale images of the faces used to create the composite.[6] Later, in 2003, Halberstadt et al reported that our attraction to averageness was also prevalent in birds, fish and even automobiles.[7]

To date, the relationship between averageness and attractiveness has not been tested for the nose specifically. We hypothesize that a digitally averaged composite nose of a specific population is rated as more attractive than the individual noses used to create it. If this proves to be so, the composite nose might serve as an aesthetic surgical reference for that specific population seeking rhinoplasty. Furthermore, using the composite nose as a reference, we attempt to calculate the degree to which an individual nasal shape deviates from average. We hypothesize that deviation away from average correlates with decreased attractiveness. If this hypothesis can be confirmed, this 'shape deviation score' could be applied to objectively quantify aesthetic rhinoplasty results.

METHODS

This pilot study was approved by the Medical Ethics Committee of the Erasmus Medical Center, Rotterdam, the Netherlands, documented by study number MEC-2015-238, and conducted between June 2016 and January 2019. Informed consent for study participation was obtained from all volunteers. Since a large number of aesthetic rhinoplasty candidates in

the Netherlands are Caucasian females, aged 25 to 40, we adhered to this specific population to test our hypotheses.

Eighty Caucasian female volunteers, who never had a (septo)rhinoplasty or extensive facial trauma, were recruited among hospital staff and patients visiting the otolaryngology outpatient clinic for complaints unrelated to the nose. For each subject, standardized frontal, oblique and profile photographs of the face in a Frankfurt horizontal plane were taken by a medical photographer. Background and lighting was similar across subjects, hair was not allowed to cover any part of the face and heavy make-up, glasses or piercings were removed. For every face image, a cropped image of the nose in isolation was created as well.

Creating the Composite Average

The first subset of 40 photographical series was used to compute an averaged composite for each view (Figure 1). A set of 40 subjects is sufficient to create a stable composite that does not significantly alter when more subjects are added. Java Psychomorph version 6 was used to create the composite (publically available at http://users.aber.ac.uk/bpt/#software, written by Dr. B. Tiddeman, department of Computer Science, Aberystwyth University, UK).[8] Standardized points were manually positioned in each face to delineate the position of facial feature landmarks of each individual image, and the average shape was calculated as the mean positions of corresponding delineated points across the face set. For each view, the images were reshaped to this average shape, and this set was then merged to a composite. A comprehensive description of this methodology, albeit for different purposes, is described elsewhere.[9]

Attractiveness Rating

A panel was organized to rate the attractiveness of the individual images of the first subset of 40 photographical series, the composite, and a second subset of 40 images. The second subset of 40 images was recruited to increase sample size, and to assess whether preference for the composite would also be evident in a comparable sub population that was different from the population the composite was created from. For each subject, including the composites, panoramically oriented images of frontal, oblique and profile views were created (Figure 2) and presented on an online interface. Each panel member was asked to provide sex, age, ethnicity, level of education and city of residence. Next, the panoramically oriented images were presented in random order in two blocks, comprising a set of 81 full faces (including the composite full face) and a set of 81 isolated noses (including the composite nose). Block

order was counterbalanced across panel members. For both blocks, the panel was asked to rate the attractiveness of the nose specifically. Attractiveness ratings were obtained using a Likert scale ranging from 1 (very ugly) to 5 (very pretty).

Calculating Deviation from Average

We hypothesize that nasal shape deviation away from average correlates with lower attractiveness ratings. For each nose, including the composite nose, 18 landmarks were systematically placed to delineate the shape of the nose in the profile view (similar to the delineation process used to create the composite). To calculate shape deviation of an original nose from the composite, the x and y positions of the 18 landmarks for each of the 80 noses were put into a principal components analysis. This procedure expresses the variation in landmark positions in a small number of components. Using the Kaiser-Guttman criterion only the components explaining the majority of variance in nose shape across the 80 faces were selected. The square root of the sum of the absolute deviations from average of these components can be used to create a 'shape deviation score' for each nose. Subsequently the correlation between this deviation score and the attractiveness score was calculated.

Statistical Analysis

Statistical analyses were performed using the IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp. A Student's independent sample *t*-test was used to compare mean age between volunteer subsets and a Student's one sample *t*-test was used to determine whether the composite ratings were statistically different from the sample mean. A Pearson correlation coefficient was calculated to indicate the correlation between shape deviation and attractiveness score.

RESULTS

The mean age of the 80 Caucasian female volunteers was 28.8 ± 3.8 years (range, 25-40 years). There was no difference in mean age between the first and second subset of volunteers (29.1 vs 28.6 years, *t*=0.52, p=0.30). The attractiveness rating panel consisted of 41 volunteers from the general public. Twenty-one were male (51%), 20 were female (49%) and the mean age of the respondents was 35.3 ± 15.6 years (range, 18-64 years). Thirty were Caucasian (73%), 3 Asian (7%) and 8 of mixed origin (20%). The level of education was fairly high with 27 volunteers possessing a bachelor's or master's degree (66%).

For each full face or isolated nose image, all raters' scores were averaged to produce a mean score. The mean score of the 81 full face images was 2.8 (range, 1.8-4.2). The nose of the composite face was rated highest, with a mean score of 4.2. The mean score of the 81 isolated nose images was 2.6 (range, 1.5-3.9). Again, the composite nose was rated highest with a score of 3.9. Both the nose in the composite face and the isolated composite nose were rated significantly higher than the distribution of ratings of the original images (t=31.24 and t=27.76 respectively, df=79, p<0.001). Hence the hypothesis that a digitally averaged composite nose of a specific population is rated as more attractive than the individual noses used to create it, cannot be rejected.

Using principal component analysis of the x and y positions of the nasal landmarks across the profile view of the 80 original noses, the first 4 components explaining 42%, 27%, 9% and 6% (a total of 84%) of the variance in profile nose shape were used to calculate the difference of each nose from the average composite. The rating of the isolated nose correlated negatively with difference from average shape (r = -0.40, n = 80, p=0.0002), implying that the larger the difference from average shape, the lower the aesthetic judgement of the nose (Figure 3). Similarly, the rating of the nose within the whole face correlated negatively with difference from average shape (r = -0.36, n = 80, p=0.00151). Therefore, the hypothesis that deviation away from average correlates with decreased attractiveness can neither be rejected.

DISCUSSION

This study shows that a composite nose, mathematically created from standardized photographs of 40 Caucasian females, is rated as more attractive than the individual noses used to create the composite. Therefore, for this specific population, an average nose is an attractive nose.

The composite nose in this study could serve as a surgical template for female Caucasians seeking aesthetic rhinoplasty. The template might aid in defining a surgeon's artistic sense of aesthetics next to the traditional rules and guidelines that are frequently used as the basis for nasofacial shape analysis and surgical planning. In this context we were interested in comparing the characteristics of the average composite 'attractive' nose to a selection of important neoclassical canons and traditional aesthetic guidelines. We found some deviations (Figure 4): The evaluation of transverse facial proportions (rule of fifths - vertical black lines) revealed an alar base that is slightly wider than the inner canthal distance. The rule of thirds in profile view (horizontal black lines; distance trichion to nasion, nasion to subnasale and subnasale to menton should be equal) does not fit as the composite has a

shorter nose and longer lower face, which is similar to the findings of Farkas' anthropological studies.[1] Nasal tip projection using the Goode method and the nasolabial angle are in line with common standards.[10,11] One must be aware that the composite found in this study only serves as a template for Caucasian female faces aged 25-40. Since the ideal nasal shape is likely not universal, the development of average composites to serve as prototypes for different rhinoplasty populations (eg, Mediterranean, Asian, African American, male) is necessary.[12]

Furthermore, we have shown that by using a principal component analysis of nasal landmarks in profile view, it is possible to calculate a shape deviation score. The more a nose deviated away from the average composite (high shape deviation score), the lower the attractiveness score of that nose was. The change of this shape deviation score following aesthetic rhinoplasty could have potential as an objective measure for the success of the procedure (Figures 5 and 6). This purpose would first necessitate further study, investigating whether shape deviation scores correlate with patient satisfaction (quantified with patient-reported outcome measures). Although patient satisfaction in the end remains the major determinant of the success of cosmetic surgery, the availability of an objective aesthetic rhinoplasty outcome measure could complement patient-reported outcome measurements by eliminating the impact of confounding factors that influence subjective ratings.

This conceptual pilot study has several limitations. Creating a composite does not only result in a nose that is average in shape, but also introduces smoothed skin texture and more symmetry. Therefore, symmetry and skin tone could have been confounding factors that contributed to the preference of the composite over the original faces. However, these confounders do not affect the demonstrated correlation between shape deviation from average and attractiveness. We showed that within the original images ($n = 80$, excluding the composite), a deviation from average score purely based on nose shape metrics and not on symmetry or skin tone is inversely related to attractiveness. These images all have original skin tone and in our opinion, there is no reason to assume that faces with a more averaged nose shape have less skin blemishes. Furthermore, the shape metric is calculated in the profile view, which is independent of nasal symmetry. As such the performed experiment supports the influence of averageness on attractiveness, independently of symmetry and skin tone. Additionally, previous research has shown that for the face as a whole the importance of averageness in the judgement of attractiveness is indeed quite robust: several authors have demonstrated that preference for averageness remains when corrected for symmetry or skin tone.[9,13-19] A second limitation is the need for manual positioning of the delineation

landmarks. Despite our standardized and meticulous placement, there is a chance of a small degree of inconsistency or error. There has been marked progress in the automated delineation of facial landmarks in the past two decades.[20-23] Although similar automation could be applied to the nose, automation reduces but does not eliminate all error. Third, in our panel of raters, the number of Caucasians was high as well as their level of education. Nevertheless, the volunteers with a lower level of education still preferred the average composite nose. Fourth, although average faces and noses are perceived as attractive, averageness is not a substitute for the aesthetic ideal. When creating a 'high average composite' of the 25 percent most attractive faces, this prototype is rated as even more attractive.[24] This implies that averageness alone does not completely explain attractiveness.

CONCLUSION

This conceptual study shows that in Caucasian females, mathematically averaged noses are attractive. Furthermore, we were able to calculate a score expressing deviation from average nasal shape, which correlated negatively with attractiveness. Change in this nasal shape deviation score following rhinoplasty has potential as an objective aesthetic rhinoplasty outcome measure, although the correlation between averageness, attractiveness and patient satisfaction needs further research.

REFERENCES

1. Farkas LG, Hreczko TA, Kolar JC, Munro IR. Vertical and horizontal proportions of the face in young adult North American Caucasians: revision of neoclassical canons. *Plast Reconstr Surg.* 1985;75(3):328-338.

2. Farkas LG, Kolar JC. Anthropometrics and art in the aesthetics of women's faces. *Clin Plast Surg.* 1987;14(4):599-616.

3. Torsello F, Mirigliani L, D'Alessio R, Deli R. Do the neoclassical canons still describe the beauty of faces? An anthropometric study on 50 Caucasian models. *Prog Orthod.* 2010;11(1):13-19.

4. Bashour M. History and current concepts in the analysis of facial attractiveness. *Plast Reconstr Surg.* 2006;118(3):741-756.

5. Rubenstein AJ, Langlois JH, Roggman LA. What makes a face attractive and why: The role of averageness in defining facial beauty. In: Rhodes G, Zebrowitz LA, eds. *Advances in Visual Cognition-Facial Attractiveness: Evolutionary, Cognitive, Cultural, and Motivational Perspectives.* Westport, Conn.: Ablex; 2001:1-33.

6. Langlois JH, Roggman LA. Attractive Faces Are Only Average. *Psychol Sci.* 1990;1(2):115-121.

7. Halberstadt J, Rhodes G. It's not just average faces that are attractive: computer-manipulated averageness makes birds, fish, and automobiles attractive. *Psychon Bull Rev.* 2003;10(1):149-156.

8. Tiddeman B, Burt DM, Perrett D. Computer Graphics in Facial Perception Research. *IEEE Computer Graphics and Applications.* 2001;21(5):42-50.

9. Valentine T, Darling S, Donnelly M. Why are average faces attractive? The effect of view and averageness on the attractiveness of female faces. *Psychon Bull Rev.* 2004;11(3):482-487.

10. Goode RL. A method of tip projection measurement. In: Powell N, Humphreys B, eds. *Proportions of the aesthetic face*: New York: Thieme-Stratton; 1984:15-39.

11. Boahene KDO, Orten SS, Hilger PA. Facial analysis of the rhinoplasty patient. In: Papel ID, Frodel JL, Holt GR, et al., eds. *Facial plastic and reconstructive surgery*. New York: Thieme; 3rd edition; 2009:477-487.

12. Broer PN, Buonocore S, Morillas A, et al. Nasal aesthetics: a cross-cultural analysis. *Plast Reconstr Surg.* 2012;130(6):843e-850e.

13. Rhodes G, Sumich A, Byatt G. Are average facial configurations attractive only because of their symmetry? *Psychol Sci.* 1999;10:52-58.

14. Rhodes G, Yoshikawa S, Clark A, Lee K, McKay R, Akamatsu S. Attractiveness of facial averageness and symmetry in non-western cultures: in search of biologically based standards of beauty. *Perception.* 2001;30(5):611-625.

15. Langlois JH, Roggman LA, Musselman L. What is average and what is not average about attractive faces? *Psychol Sci.* 1994;5:214.

16. Jones BC, DeBruine LM, Little AC. The role of symmetry in attraction to average faces. *Percept Psychophys.* 2007;69(8):1273-1277.

17. Rhodes G, Tremewan T. Averageness, exaggeration, and facial attractiveness. *Psychol Sci.* 1996;7:105-110.

18. Little AC, Hancock PJ. The role of masculinity and distinctiveness in judgments of human male facial attractiveness. *Br J Psychol.* 2002;93(Pt 4):451-464.

19. O'Toole AJ, Price T, Vetter T, Bartlett JC, Blanz V. 3D shape and 2D surface textures of human faces: The role of "averages" in attractiveness and age. *Image & Vision Computing.* 1999;18:9-19.

20. Haghighat M, Abdel-Mottaleb M, Alhalabi W. Fully automatic face normalization and single sample face recognition in unconstrained environments. *Expert Systems with Applications.* 2016;47:23-34.

21. Lanitis A, Taylor CJ, Cootes TF. Automatic interpretation and coding of face images using flexible models. *IEEE Transactions on Pattern Analysis and machine intelligence.* 1997;19(7):743-756.

22. Fabian Benitez-Quiroz C, Srinivasan R, Martinez AM. Emotionet: An accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 2016:5562-5570.

23. Zulqarnian Gilani S, Shafait F, Mian A. Shape-based automatic detection of a large number of 3D facial landmarks. . *In Proceedings of the IEEE conference on computer vision and pattern recognition* 2015:4639-4648.

24. Perrett DI, May KA, Yoshikawa S. Facial shape and judgements of female attractiveness. *Nature.* 1994;368(6468):239-242.

Figure Legend

Figure 1. Digitally averaged composite of 40 images of randomly selected Caucasian females aged 25-40 years (frontal view example).

Figure 2. Examples of panoramically oriented images as presented to the panel of raters. Example of original full face of a 28-year-old female subject: (A) profile left, (C) oblique left, (E) frontal, (G) oblique right, (I) profile right; example of original isolated nose of same subject, (K) profile left, (M) oblique left, (O) frontal, (Q) oblique right, (S) profile right; composite full face, (B) profile left, (D) oblique left, (F) frontal, (H) oblique right, (J) profile right; and composite isolated nose, (L) profile left, (N) oblique left, (P) frontal, (R) oblique right, and (T) profile right.

Figure 3. The rating of the nose across the faces (yellow circles) correlated negatively with difference from average estimate (r=-0.40, n=80, p=0.0002); the larger the difference score the lower the aesthetic judgment of the nose shape. The thin red line represents the best fit regression line relating nose rating to the difference from average score. X denotes the rating of the composite nose from the blend of 40 faces. The shape difference of the composite nose is not zero since the composite was computed from the first set of 40 faces, not all 80 faces.

Figure 4. A selection of ratios, angles and proportions of the composite nose. Horizontal black lines: rule of thirds. Vertical black lines: rule of fifths. Blue: nasolabial angle, 104°. Red: nasal tip projection according to Goode, 0.57. Yellow: nasofrontal angle, 139°.

Figure 5. (A) Example of original nose (32-year-old female subject) (B) reshaped by adjusting the landmark position to that of the averaged composite.

Figure 6. (A) Example of original nose (34-year-old female subject) (B) reshaped by adjusting the landmark position to that of the averaged composite.
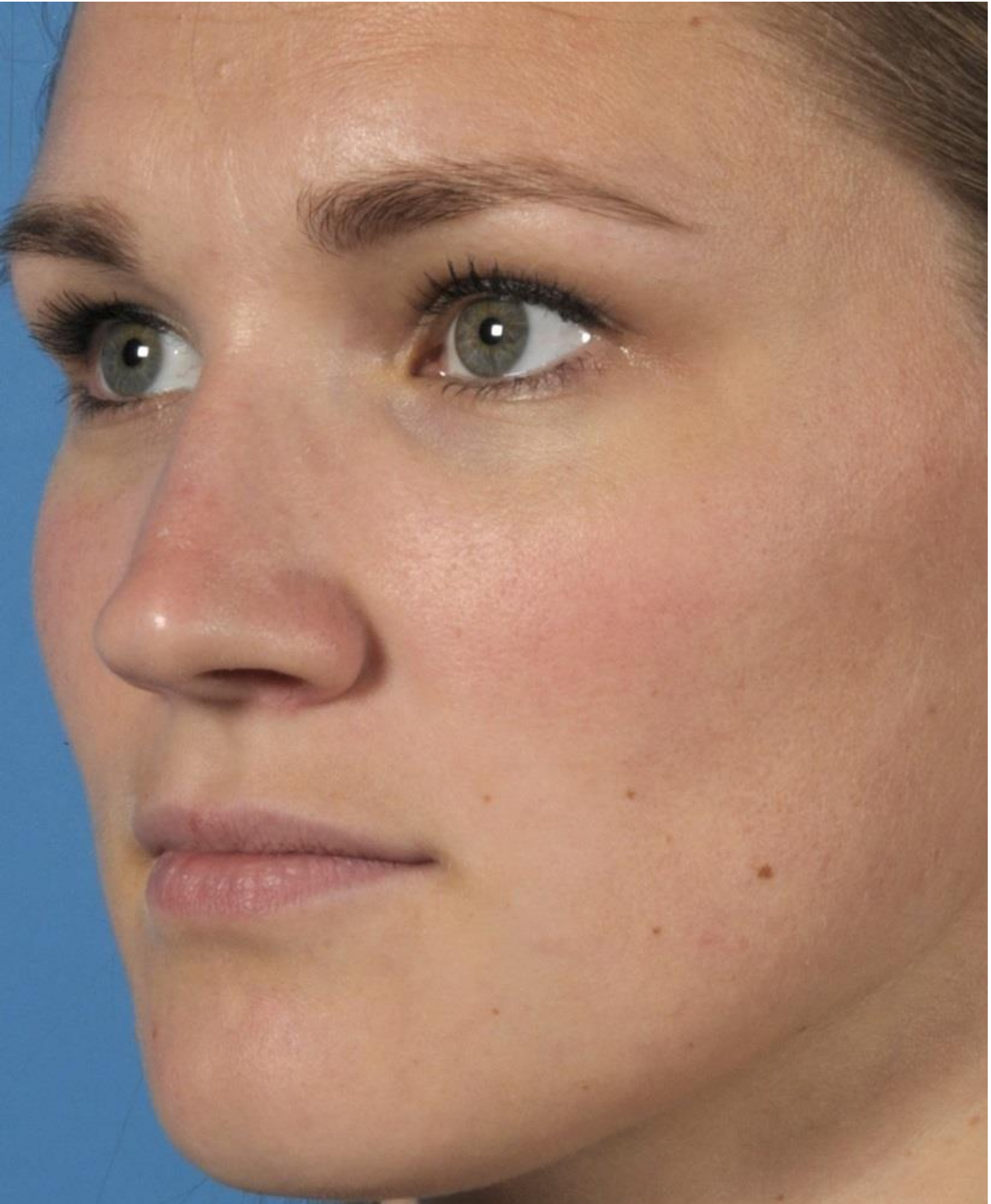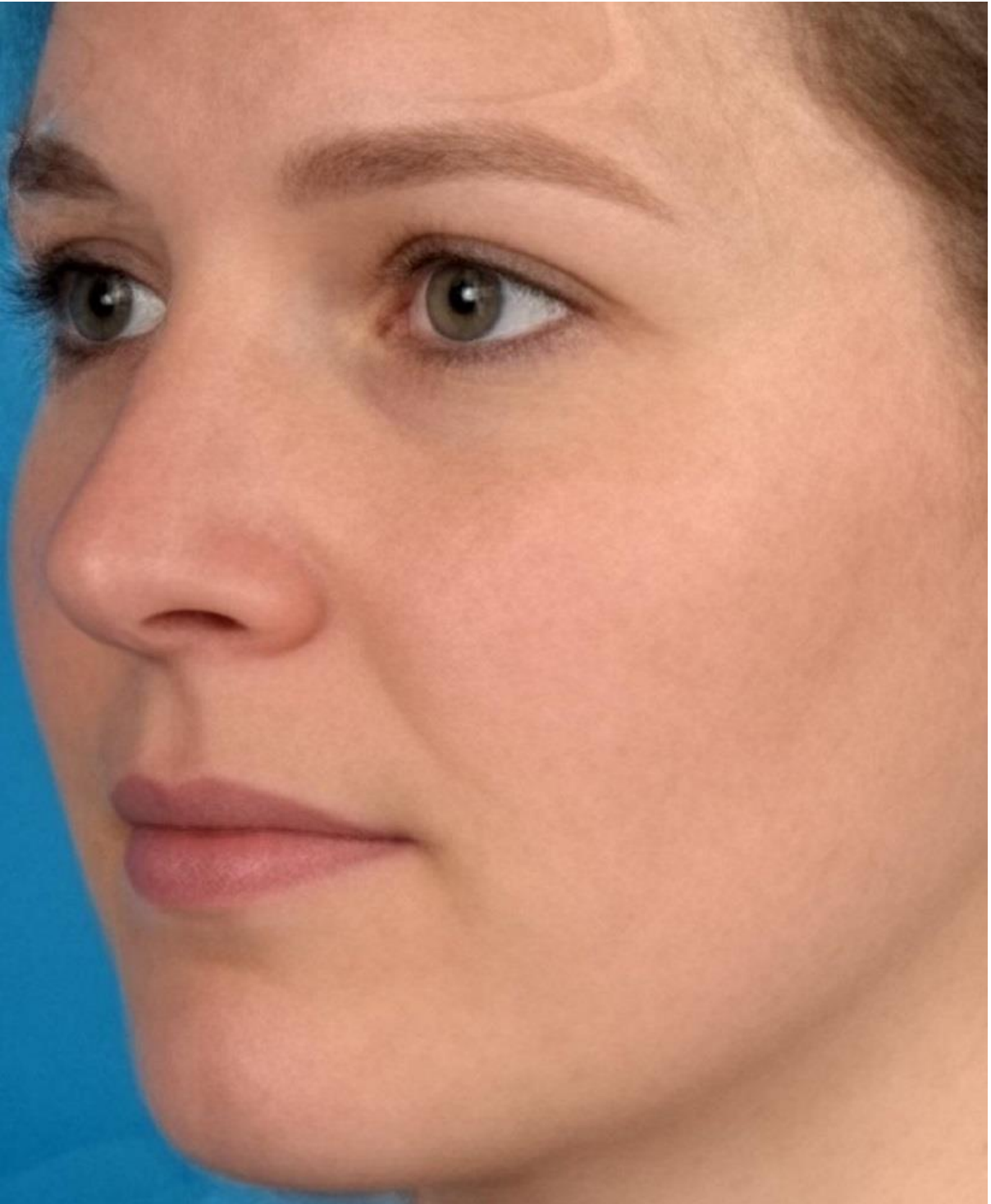
Figure_1

Figure_2A
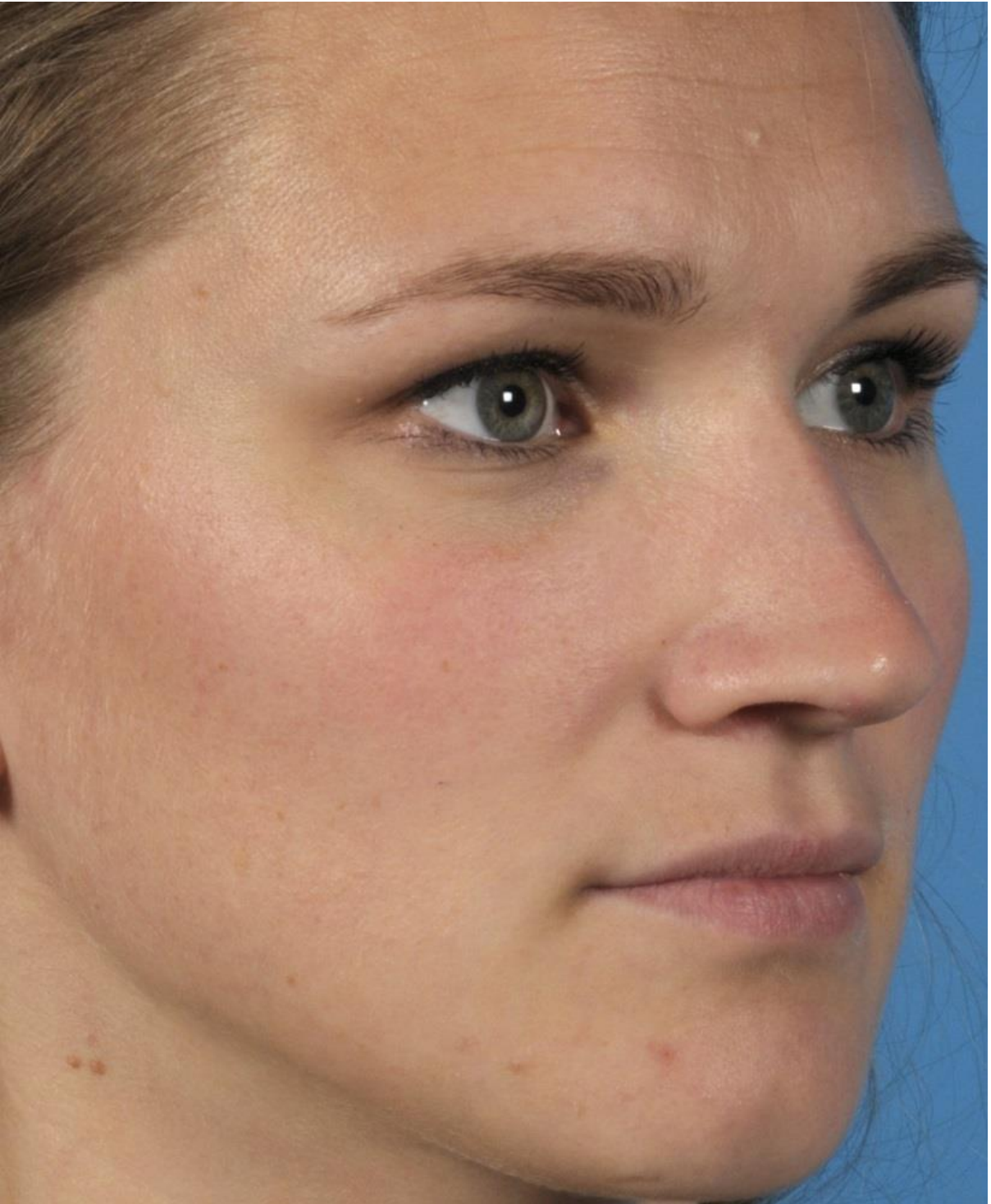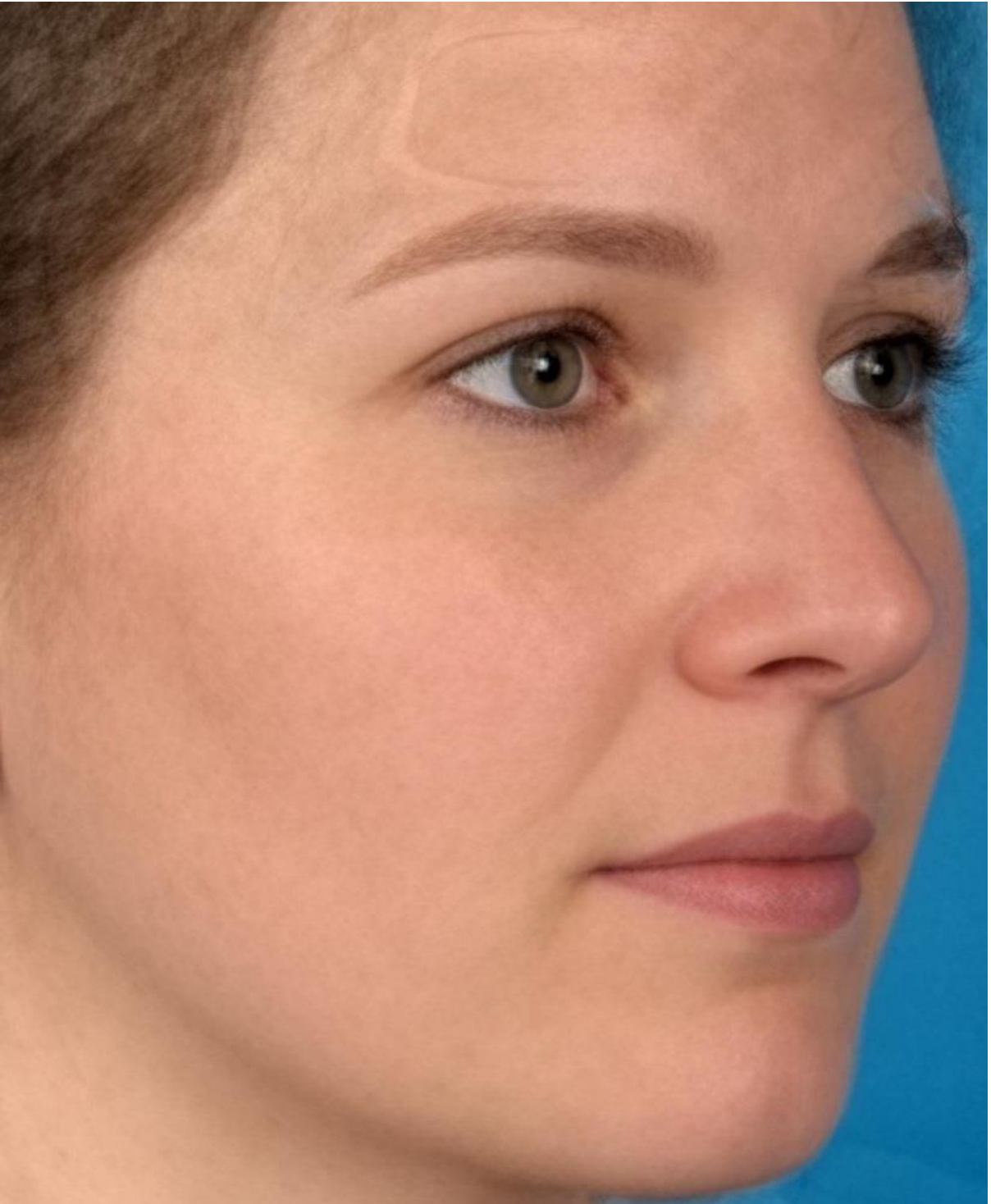
Figure_2B

Figure_2C

Figure_2D

Figure_2E

Figure_2F

Figure_2G

Figure_2H

Figure_2I

Figure_2J
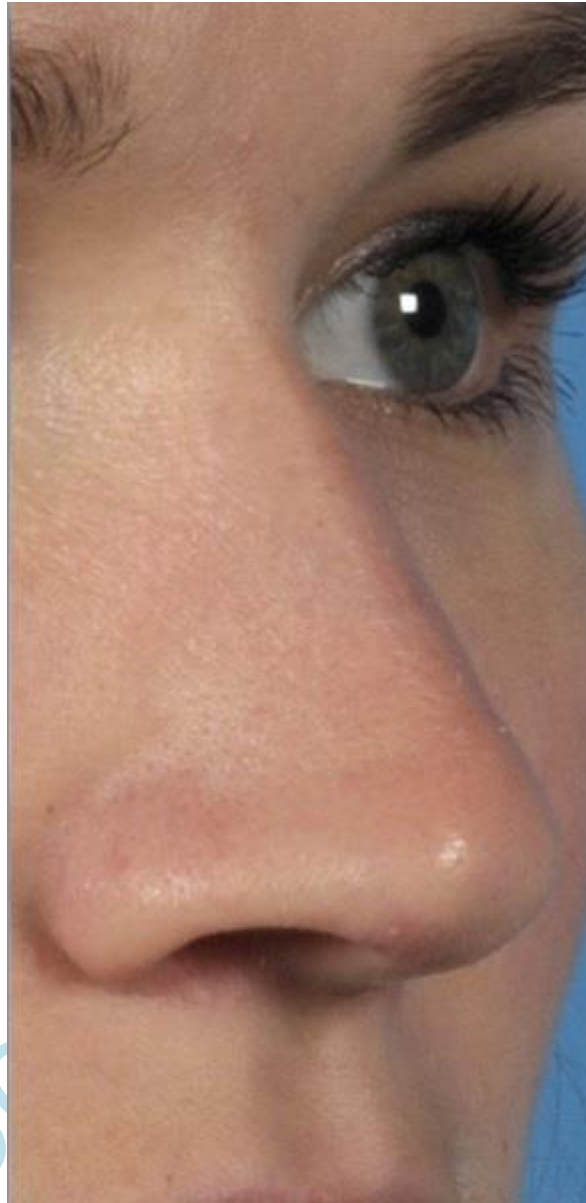
Figure_2K
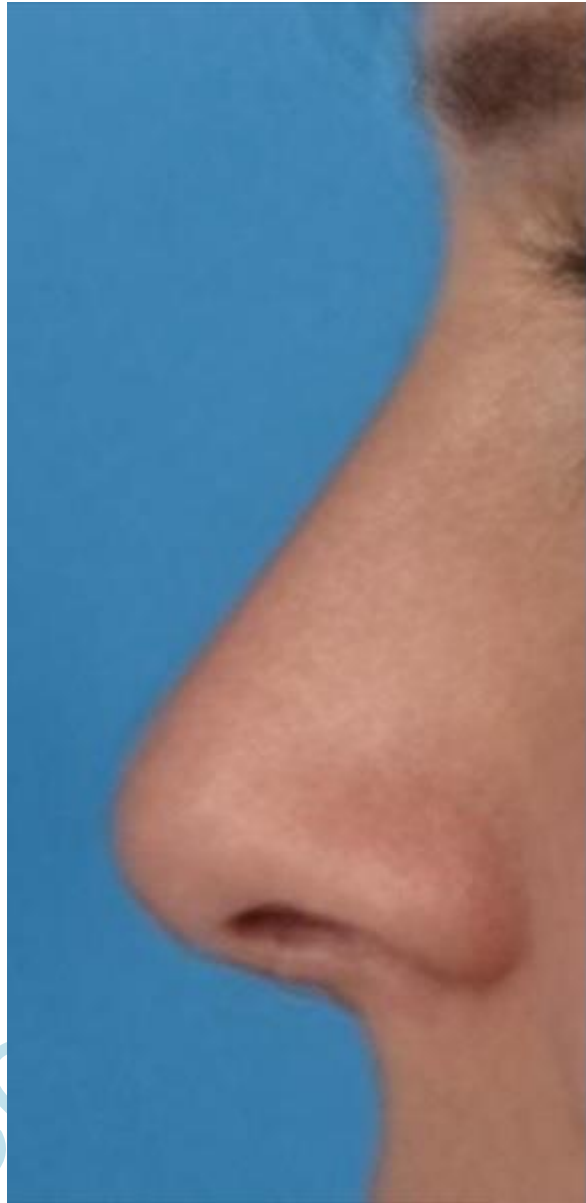
Figure_2L

Figure_2M

Figure_2N

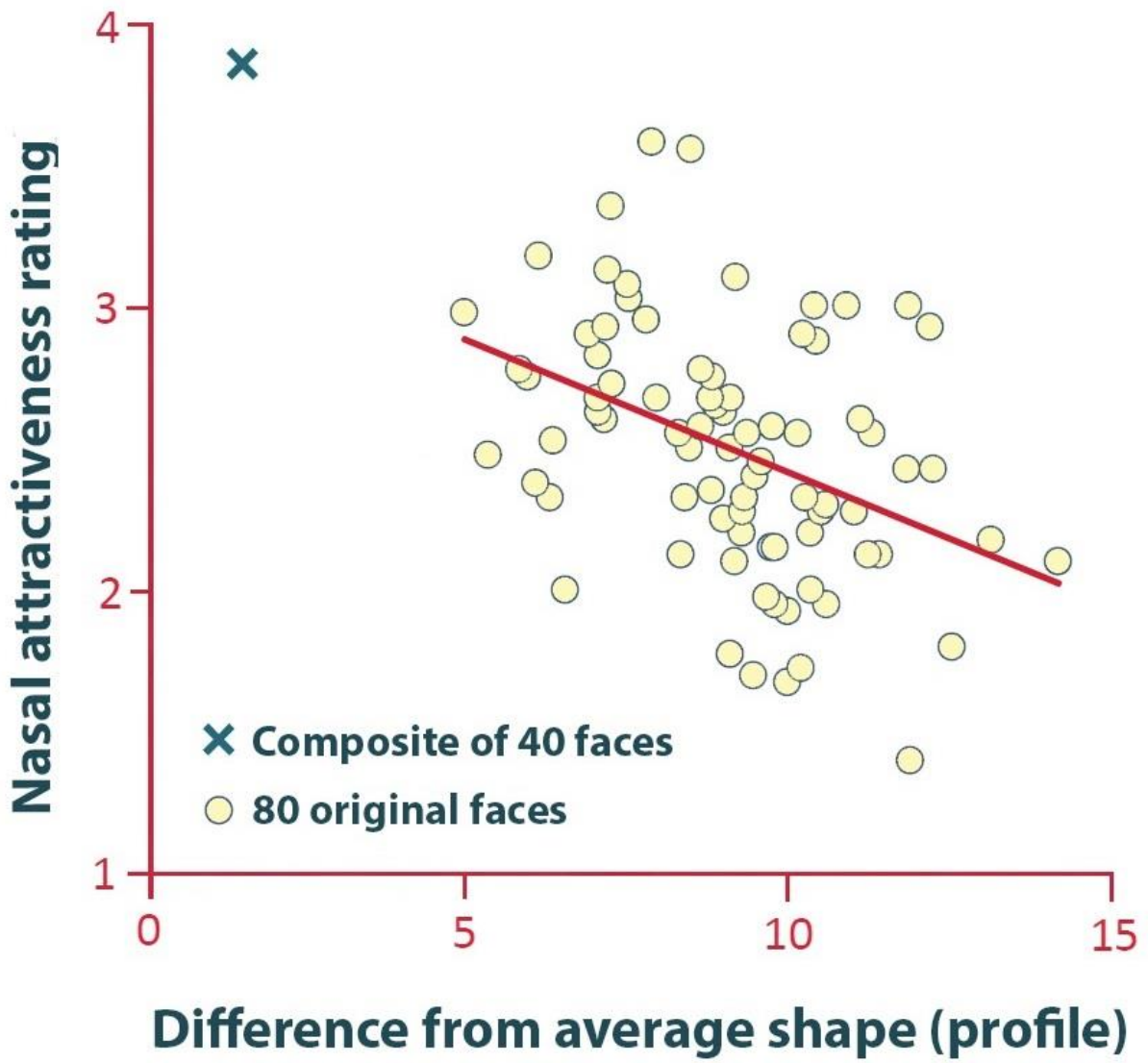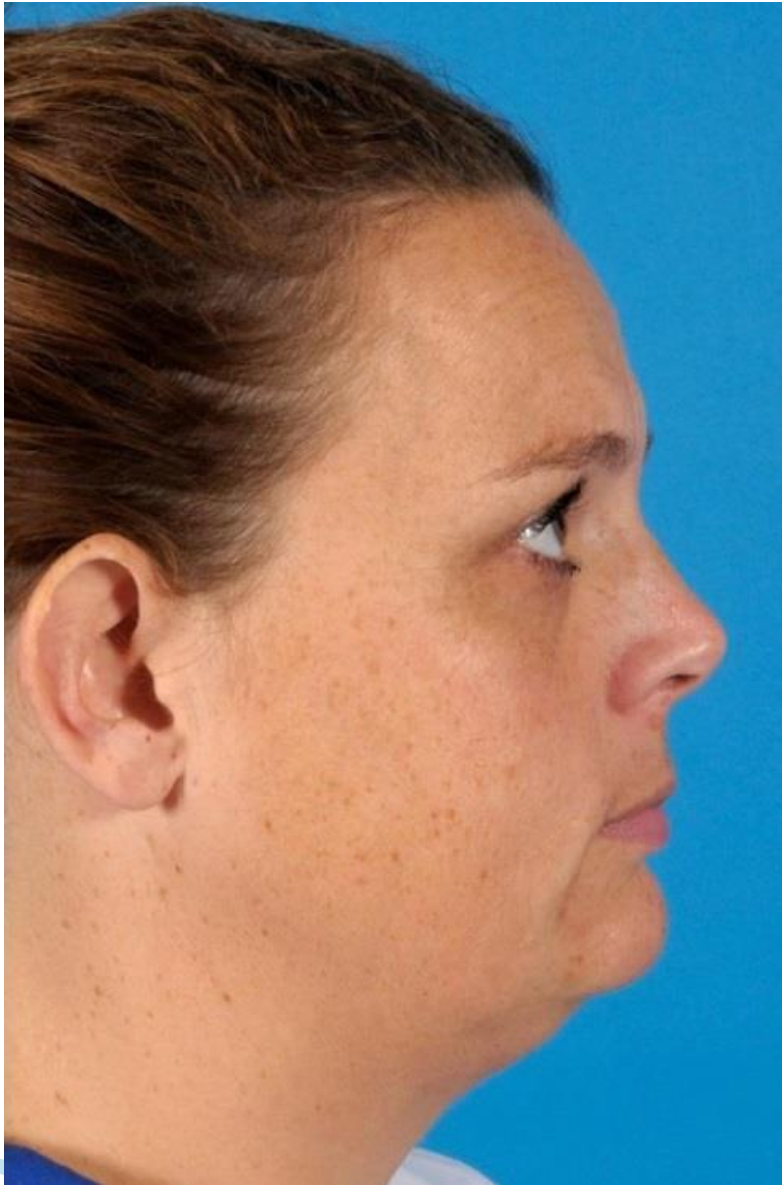Figure_2O

Figure_2P

Figure_2Q

Figure_2R

Figure_2S

Figure_2T

Figure_3

Figure_4

Figure_5A

Figure_5B

Figure_6A

Figure_6B