Interrater reliability of a Phenotypic Assessment Tool for the Ear Morphology in Microtia

Daniela V Luquetti[1,2,3], **Babette S Saltzman**[2], **Kathleen C Sie**[2,4], **Craig B Birgfeld**[2,5], **Brian G Leroux**[6], **Kelly N Evans**[1,2], **James M Smartt Jr**[2,5], **David D Tieu**[2,4], **Daniel J Dudley**[2], and **Carrie L Heike**[1,2]

[1]Department of Pediatrics, University of Washington, Seattle, WA

[2]Craniofacial Center, Seattle Children's Hospital, Seattle, WA

[3]Center of Tissue and Cell Sciences, Seattle Children's Research Institute, Seattle, WA

[4]Division of Pediatric Otolaryngology, Head and Neck Surgery, Seattle Children's Hospital, Seattle, WA

[5]Division of Plastic Surgery, Department of Surgery, University of Washington, Seattle, WA

[6]Departments of Biostatistics, School of Public Health and Oral Health Sciences, School of Dentistry, University of Washington, Seattle, WA

Abstract

The Elements of Morphology Standard Terminology working group published standardized definitions for external ear morphology. The primary objective of our study was to use these descriptions to evaluate the interrater reliability for specific features associated with microtia. We invited six raters from three different subspecialties to rate 100 ear photographs on 32 features. We calculated overall and within specialty and professional experience intraclass correlation coefficients (ICC) and 95% confidence intervals. A total of 600 possible observations were recorded for each feature. The overall interrater reliability ranged from 0.04 (95%CI: 0.00–0.14) for the width of the antihelix inferior crus to 0.93 (95%CI: 0.91–0.95) for the presence of the inferior crux of the antihelix. The reliability for quantitative characteristics such as length or width of an ear structure was generally lower than the reliability for qualitative characteristics (e.g. presence or absence of an ear structure). Categories with very poor interrater reliability included anti-helix inferior crux width (0.04, 95%CI: 0.00–0.14), crux helix extension (0.17, 95% CI 0.00–0.37), and shape of the incisura (0.14, 95%CI: 0.01–0.27). There were no significant differences in reliability estimates by specialty or professional experience for most variables. Our study showed that it is feasible to systematically characterize many of structures of the ear that are affected in microtia. We incorporated these descriptions into a standardized phenotypic assessment tool (PAT-Microtia) that might be used in multicenter research studies to identify sub-phenotypes for future studies of microtia.

CORRESPONDING AUTHOR: Daniela V Luquetti, 1900 9th Avenue, Room 520, Seattle, WA 98101. daniela.luquetti@seattlechildrens.org, Phone: (206)-884-5120, Fax: (206)-884-1405.

INTRODUCTION

The complexity of characterizing variation in external ear anatomy presents a challenge to clinicians and researchers. Classification systems by Marx [1926], Weerda [1988], and Tanzer [1978] have been created to describe the severity and types of ear anomalies. However, these systems use broad classifications that do not capture all of the complexities of ear malformations. In 2009, the Elements of Morphology: Standard Terminology (EMST09) working group published a series of articles to standardize the terms used to describe human morphology and congenital anomalies [Allanson et al., 2009]. One article provided detailed definitions and illustrations developed by a panel of experts, the Ear Dysmorphology Subgroup (EDS), of the normal anatomy and variations of the external ear [Hunter et al., 2009a]. This valuable reference is available at the National Human Genome Research Institute (http://elementsofmorphology.nih.gov/index.cgi?lid=0063bc07cdb40b0b). Hunter et al [2009b] demonstrated the feasibility of using such descriptions to better characterize the external ear in a study that compared characteristics of ears in the general population to the ears from patients with Cornelia de Lange syndrome (CdLS).

Molecular genetics studies, such as genome-wide association and exome sequencing studies have demonstrated that poor phenotypic characterization of a sample population can impair the study results [Ng et al., 2009]. These findings support the need for careful, standardized phenotypic classification of anomalies in medical genetic studies. Pooling cases with broadly similar congenital anomalies could preclude our ability to recognize subphenotypes that arise from different genetic causes. For example, different genetic risk factors in one cohort of children with clefts were only identified once phenotypic groups with (1) isolated cleft lip and (2) isolated cleft lip and palate were distinguished in the data analysis [Jugessur et al., 2010]. Development of a classification tool that enables the identification of subphenotypes among individuals with microtia may help researchers to identify a genetic basis for these common anomalies.

The primary objectives of this study are to evaluate the interreliability of the standardized descriptions by the EDS and a modified version of the categories used by Hunter in his study in CdLS [Hunter et al., 2009a; Hunter et al., 2009b] for patients with microtia. We also sought to create a phenotypic assessment tool that incorporated these descriptions and could be used to identify sub-phenotypes for future studies of microtia.

METHODS

Sample selection

We identified individuals with a diagnostic code developed by the Craniofacial Clinic for microtia or craniofacial microsomia using the Craniofacial Clinic database at Seattle Children's Hospital as part of a larger study. All participants were under 21 years of age at the time of ascertainment. Medical charts were reviewed in a random order and individuals who had a confirmed diagnosis of microtia were included, regardless of whether the individuals' microtia was isolated, associated with multiple anomalies, or part of a syndrome. Small ears with no abnormalities of any ear structure were not included in this study. Clinical photographs obtained by a trained medical photographer are part of the standard care in the Craniofacial Clinic. We reviewed images in the clinic photographic repository for eligible individuals and selected those with photographs of the lateral view that allowed for adequate visualization of the ear.

We selected the first 84 patients who met the inclusion criteria and had photographs of adequate quality. The photographer selected the optimal image for individuals with more

than one photograph. Both sides were used as independent photographs for the two individuals with bilateral microtia. All photographs were cropped and oriented in a left profile view. Fourteen unaffected contralateral ears from individuals with unilateral microtia were included as controls. Thus, the final dataset of 100 ear photographs from 84 individuals included 86 ears with microtia and 14 unaffected ears.

Raters' selection

We invited six raters from the following three medical specialties: pediatrics, plastic surgery, and otolaryngology All providers had a minimum of one year of clinical training in the diagnosis and management of children with microtia, and all were members of the Craniofacial Center at our institution at the time of this study.. We chose one more experienced and one less experienced professional (fellow or junior faculty) from each specialty. Raters had not previously agreed on standardized descriptions for microtia, nor had they used the EDS terms systematically in clinical practice. Raters were blinded from each other's ratings. We recorded data on years since fellowship, experience evaluating patients with microtia, and total time spent rating.

Development of the standardized reporting system

We selected the categories utilized by Hunter et al [2009b] that appeared to be most relevant for features altered in microtia, with minor modifications (Table III). We provided raters with an Excel spreadsheet containing these fields along side the study ear photograph. Some rating categories were qualitative, such as "absence and presence", while others were quantitative, such as "folded, partially-folded, over-folded". Raters were asked to choose values from a dropdown menu for each component, ranging from normal to absent (example: normal, underdeveloped, absent). We added four different options for 'unable to rate", which included: (1) unable to rate due to photo (e.g.: hair in front of the ear), (2) unable because feature was absent (e.g.: absent helix, thus could not rate "shape of helix"), (3) unable for another reason (e.g.: reconstructive surgery of the ear), and (4) missing data (i.e.: category left blank by the rater). We asked raters to provide details if they chose "unable to rate for other reason" and to record any other free-text comments. No measurement scale was provided in the photograph. (See scoring sheet at Appendix 1 –see supporting information online).

Written and pictorial anatomic definitions based on the consensus published by the EDS [Hunter et al., 2009a] were provided to raters along with specific instructions about the scoring system. We did not provide formal training on the definitions or the reporting system. The scoring sheets were sent by a secure web server. All raters were asked to rate the photographs in the same order. There were two rating sessions of 50 photographs each; a two-month period elapsed between sessions. We held an in-person meeting with raters between the first and the second session to discuss their impressions, concerns and potential improvements. We specifically asked about ease of use of the rating system and clarity of instructions. We incorporated minor modifications to the tool and instructions based on raters' suggestions from the first session, and added the following categories: overall shape of the helix, shape of superior helix, and concha morphology. We only incorporated those features with a reasonable reliability into the phenotypic assessment tool for microtia.

Data analysis

We provide descriptive statistics for each ear feature (e.g.: helix, stem of the antihelix) on the prevalence, ratings, and the percentage of "unable to rate". Estimates of the prevalence for each ear feature were calculated using ratings assigned by all 6 raters (600 total ratings), since no rater could be considered the "gold standard".

We calculated the interrater reliability using the intraclass correlation coefficient (ICC) and its 95% confidence interval using the one-way analysis of variance (ANOVA) method for each individual feature [Bewick et al., 2004]. ICCs estimate the strength of the correlation among ratings on the same ear by different raters and the calculation accounts for systematic and random errors. An ICC of one indicates perfect correlation and zero indicates no correlation. We calculated ICCs for all 6 raters as a group and we compared ICC values by (1) experience level and (2) discipline of the raters using an informal approach by which two ICC values were declared different if their 95% confidence intervals did not overlap. We also performed separate analyses on the two groups of 50 images to check for changes in reliability between the rating sessions.

In a secondary analysis, we sought to evaluate the reliability of a post hoc, simplified composite rating. For this specific assessment, we classified each ear feature as normal or abnormal. We considered a specific feature to be "normal" when the feature (eg.: antihelix stem) was present and presented the average dimension or shape and "abnormal" when absent or above or below the average category in length, width, or shape (as applicable) was reported. Thus, abnormal was used only as a term of convenience and for this specific analysis. We calculated the ICC using these collapsed categories.

RESULTS

Rater's characteristics

The less experienced raters had between one and four years of experience evaluating patients with microtia and the more experienced raters had at least five years of experience. The reported total time spent for rating ranged from 2 to 2.75 hours between raters (Table I).

Characteristics of the Sample

The patients were mostly males (70.1%) and their mean age was 6.1 years (standard deviation: 4.4; range 2 weeks to 17 years). There were 100 photographs each scored by 6 raters, for a total of 600 possible observations for each of the 32 variables. The "absence of a feature" was the most common explanation for the majority of unrated features. Ears that were unanimously categorized as having more severe degrees of microtia (Marx categories III and IV) compared to milder degrees of microtia (Marx categories I and II) were more likely to have "absence of feature" selected. In those with milder degrees of microtia, the following five features had more than 60% not rated due to the "absence of the feature": superior crus width, prominence, and crus helix length, prominence and extension. In comparison, 16 of the features among the more severely microtia group were not rated over 90% of the time due to absence of the features (data not shown). Most features had less than 10% missing data (Table II). Some raters were more likely than others to indicate that they were unable to rate a feature. The proportional differences between use of the "unable to rate" categories were not correlated with speciality, nor to experience level, indicating that a rater's wiliness to use this designation was primarily determined by personal characteristics of that individual rater (data not shown).

Raters assigned the following Marx classifications to the data set: normal (18.3%), grade I (18.5%), grade II (15.1%), grade III (41.4%), and grade IV (6.6%). The frequency of specific ratings for each component of the ear is reported in Table III. These frequencies represent the proportion of values assigned to each of the 100 rated photographs by all raters for a total of 600 possible ratings of each ear component and are shown as the combined information of all raters.

Interrater reliability

The interrater reliability of the combined group of raters for the ear components ranged from 0.04 (95%CI: 0.00–0.14) for the width of the antihelix inferior crus to 0.93 (95%CI: 0.91–0.95) for the presence/absence of the inferior crux of the antihelix. The reliability for the assessment of characteristics measured quantitatively, such as length or width of a structure was generally lower than the reliability for assessing qualitative characteristics, such as presence or absence of an ear structure. The interrater reliability for qualitative characteristics ranged from 0.38 to 0.93, while for quantitative characteristics it ranged from 0.04 to 0.63. Some categories had very poor interrater reliability, including anti-helix inferior crux width (0.04, 95%CI: 0.00–0.14), crux helix extension (0.17, 95% CI 0.00–0.37), and shape of the incisura (0.14, 95%CI: 0.01–0.27). (Table IV)

We did not identify meaningful differences in reliability by specialty for most variables. However, pediatricians and otolaryngologists had better reliability than plastic surgeons for assessment of serpiginous quality of the antihelix stem (0.59 and 0.54 vs 0, respectively). In addition, pediatricians had higher reliability than plastic surgeons for assessment of the size of the tragus (0.58 versus 0.07, respectively) and higher reliability than both of the other groups for the presence/absence of the pit (0.68 versus 0.26 and 0.29, respectively). Conversely, the plastic surgeons had higher reliability than the pediatricians for the prominence of the inferior crus of the antihelix (0.53 vs 0.01, respectively) although the confidence intervals overlapped in this comparison. Both plastic surgeons and pediatricians had higher reliability than otolaryngologists for the width of the incisura (0.80 and 0.72 vs 0.23, respectively; Table IV). More experienced professionals had significantly higher reliability than those with less experience for six of the ear features (anti-helix stem-length, incisura length, lobe size, pretragal ectopia, concha, and Marx classification; data not shown). We did not observe a difference or trend in reliability between the two different rating sessions. The reliability for the composite classification of each feature type (normal or abnormal) was lower than for the presence/absence ratings for each feature and the reliability was lower for each feature when we looked within groups defined by severity of microtia (grades I & II vs. grades III & IV (data not shown).

DISCUSSION

Detailed descriptions of the ear are relevant for genetic studies of ear malformations, as well as clinical practice and outcomes assessment. Previously published classification systems and the EDS initiative on standardizing the description of the ear morphology have documented the need for using standardized descriptions of ear morphology. However, such classification systems have been variably adopted by healthcare providers who treat patients with microtia, as well as in research studies. A limitation of the existing classifications systems is that they do not capture the morphology of the specific components of the ear. Our study showed that it is feasible to characterize the different parts of the ear in a systematic way and, for a proportion of the ear features, in a reliable way. Most importantly, it shows that the data captured by this type of characterization can be used to assess presence or absence of specific ear malformations as well as help to identify sub-phenotypes of ear malformations.

We observed high reliability estimates for qualitative features and lower reliability for characteristics that are expressed as a continuum. Nonetheless, some of these quantitative features can be assessed with reasonable reliability, and their use may be justified. Rating of some features presented very low reliability in individuals with microtia, such as width of the inferior crus of the antihelix, extension of crus helix, and shape of the incisura and their use must be considered carefully. Alternatively, specific training for classification of those features may improve the reliability estimates.

A wide range in phenotypic variation among ears with the same degree of microtia was observed (Fig 1). In addition, severely affected ears often lack most (or all) recognizable features, which limits the ability to capture phenotypic variation with the existing EDS ear morphology categories. This underscores the importance of obtaining photographs that could be attached to the rating form and available for re-evaluation in the future, when we might have more information to characterize that ear phenotype.

We estimated that the raters took approximately 1.5 minutes to rate each ear based on the total time spent by raters. Though this appears to be a reasonable amount of time to be spent in the clinical or research setting, this may be an underestimate for future studies with a higher proportion of ears with milder phenotypes and raters with less experience treating individuals with microtia.

The reliability for our composite simplified classification (normal/abnormal) for each feature was lower than that for the individual features. Therefore, although it is tempting to simplify a classification system by having fewer categories, our study suggests that detailed categories are required for clinical and research uses in which it is important to capture complex ear morphology.

One of the strengths in our study is that it incorporated the definitions and photographs selected by an independent, internationally recognized panel of dysmorphologists and medical geneticists (Dysmorphology Working Group-Ear Dysmorphology Subgroup). In addition, we used clinical photographs from a large craniofacial center, which allowed us to select high quality images from a large number of patients with wide variability in the severity of their ear abnormalities. Including six raters with different training and experience allowed us to evaluate the interrater reliability among all raters, as well as the effect that the variation in medical training and experience might have on reliability. Photographs were taken by a trained medical photographer, which decreased the chances of having ear components unrated because of photograph quality. Scoring sheets were promptly reviewed and returned to raters if there were un-scored variables ensuring a nearly complete data set.

There are also some limitations to our study. Close up photographs of the ear did not allow for assessment of relative ear position. In addition, raters could not determine the actual size of the ear using photographs that did not include a measurement scale. However, with the exception of ear height and width, it would not have been possible to compare those measurements with "normal" because currently there is no normative data on individual ear components. In the absence of quantifiable measures for most components of the ear, it remains important to understand the reliability of currently available tools for evaluating patients with microtia. We were unable to assess the reliability within speciality and experience categories while simultaneously adjusting for the effects of the other due to having only two raters of each speciality, each with different levels of experience. Therefore, the within-speciality reliability reported here is likely, in part, a reflection of the individuals' level of experience and vice versa. Finally, it is possible that the reliability of the quantitative traits would improve with training, particularly for those with very low reliability such as the incisura shape. While we did not administer formal training, we did provide expert raters with definitions and detailed instructions. Our study results suggest that this tool could be used by healthcare providers without formal training which would facilitate use of such a tool in a research or clinical setting. Future studies to assess the degree of improvement in reliability estimates after formal training would be helpful and the rating tool could be further tailored to address specific study hypotheses.

Standardized Reporting System for Ear Morphology

As a result of this study we have created a standardized reporting system that we believe will be very useful in the characterization and reporting of ear morphology both in the clinical setting and in multicenter research studies. This phenotypic assessment tool for microtia (PAT-Microtia) (Fig 2) includes an illustration of a typical ear, along with detailed descriptions of the various components of the ear, enabling the classification of each component in relation to its presence and the kind of abnormality it presents along with definitions and/or photo examples. We recommend attaching the ear photograph to the reporting form when possible for re-examination.

This tool can be used for the study of cases of microtia with unknown etiology and enable clustering of similar phenotypes based on the whole ear or abnormalities of a specific part of the ear. For example, such a system could be used to distinguish among cases that have absence of the helix independent of other ear abnormalities and cases that share maximum similarity of the ear. These similarities may not be related to the severity of microtia, but instead, to specific phenotypic characteristics of the ear morphology. We note that this tool might be more useful for less severe phenotypes, such as grades I and II than to more severe ones like grade III or anotia, where few, if any, features of the ear can be characterized. We hypothesize that different types of microtia have different etiologies, and use of this tool to group patients based on their refined ear phenotype may facilitate identification of different etiologies of microtia and other ear abnormalities.

The standardized assessment of ear morphology might also allow for a refined characterization of the ear among patients with syndromes. This refined characterization may help clinicians more accurately diagnose syndromes based on shared characteristics. This may be useful in syndromes thought not to have characteristic changes in ear morphology, such as CdLS. As an example, Hunter et al [2009b] found preliminary evidence that the ears of individuals with this syndrome do differ from those of the general population and that many of those differences were highly statistically significant.

A high quality photograph is necessary in order to optimize the reliability of this rating tool. We have described in a previous study [Heike et al., 2011] the best practices for photographs of the face, including the ear, in individuals with microtia using a standard camera. The use of a paper scale taped adjacent to the ear when photographing is recommended in order to record the actual size of the ear and its components

Summary

The use of a standardized reporting system for ear morphology will enable clinicians and researchers to group individuals by their specific ear phenotypes. This information might be used for a myriad of studies, both genetic and non-genetic, as well as planning of surgical interventions and evaluation of clinical outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

Allanson JE, Biesecker LG, Carey JC, Hennekam RC. Elements of morphology: introduction. Am J Med Genet A. 2009; 149A:2–5. [PubMed: 19127575]

Bewick V, Cheek L, Ball J. Statistics review 9: one-way analysis of variance. Crit Care. 2004; 8:130–136. [PubMed: 15025774]

Heike CL, Stueckle LP, Stuhaug ET, Pimenta LA, Drake AF, Vivaldi D, Sie KC, Birgfeld CB. Photographic protocol for image acquisition in craniofacial microsomia. Head Face Med. 2011; 7:25. [PubMed: 22208766]

Hunter A, Frias JL, Gillessen-Kaesbach G, Hughes H, Jones KL, Wilson L. Elements of morphology: standard terminology for the ear. Am J Med Genet A. 2009a; 149A:40–60. [PubMed: 19152421]

Hunter AG, Collins JS, Deardorff MA, Krantz ID. Detailed assessment of the ear in Cornelia de Lange syndrome: comparison with a control sample using the new dysmorphology guidelines. Am J Med Genet A. 2009b; 149A:2181–2192. [PubMed: 19764039]

Jugessur A, Shi M, Gjessing HK, Lie RT, Wilcox AJ, Weinberg CR, Christensen K, Boyles AL, Daack-Hirsch S, Nguyen TT, Christiansen L, Lidral AC, Murray JC. Fetal genetic risk of isolated cleft lip only versus isolated cleft lip and palate: A subphenotype analysis using two population-based studies of orofacial clefts in scandinavia. Birth Defects Res A Clin Mol Teratol. 2010

Marx, H. Die Missbildungen des ohres. In: DAK, O., editor. Handbuch der Spez Path Anatomie Histologie. Berlin, Germany: Springer; 1926. p. 131

National Human Genome Research Institute. Standard Terminology for the Ear. 2012. http://elementsofmorphology.nih.gov/index.cgi?lid=0063bc07cdb40b0b

Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, Shaffer T, Wong M, Bhattacharjee A, Eichler EE, Bamshad M, Nickerson DA, Shendure J. Targeted capture and massively parallel sequencing of 12 human exomes. Nature. 2009; 461:272–276. [PubMed: 19684571]

Tanzer RC. Microtia. Clin Plast Surg. 1978; 5:317–336. [PubMed: 359217]

Weerda H. Classification of congenital deformities of the auricle. Facial Plast Surg. 1988; 5:385–388. [PubMed: 3270622]
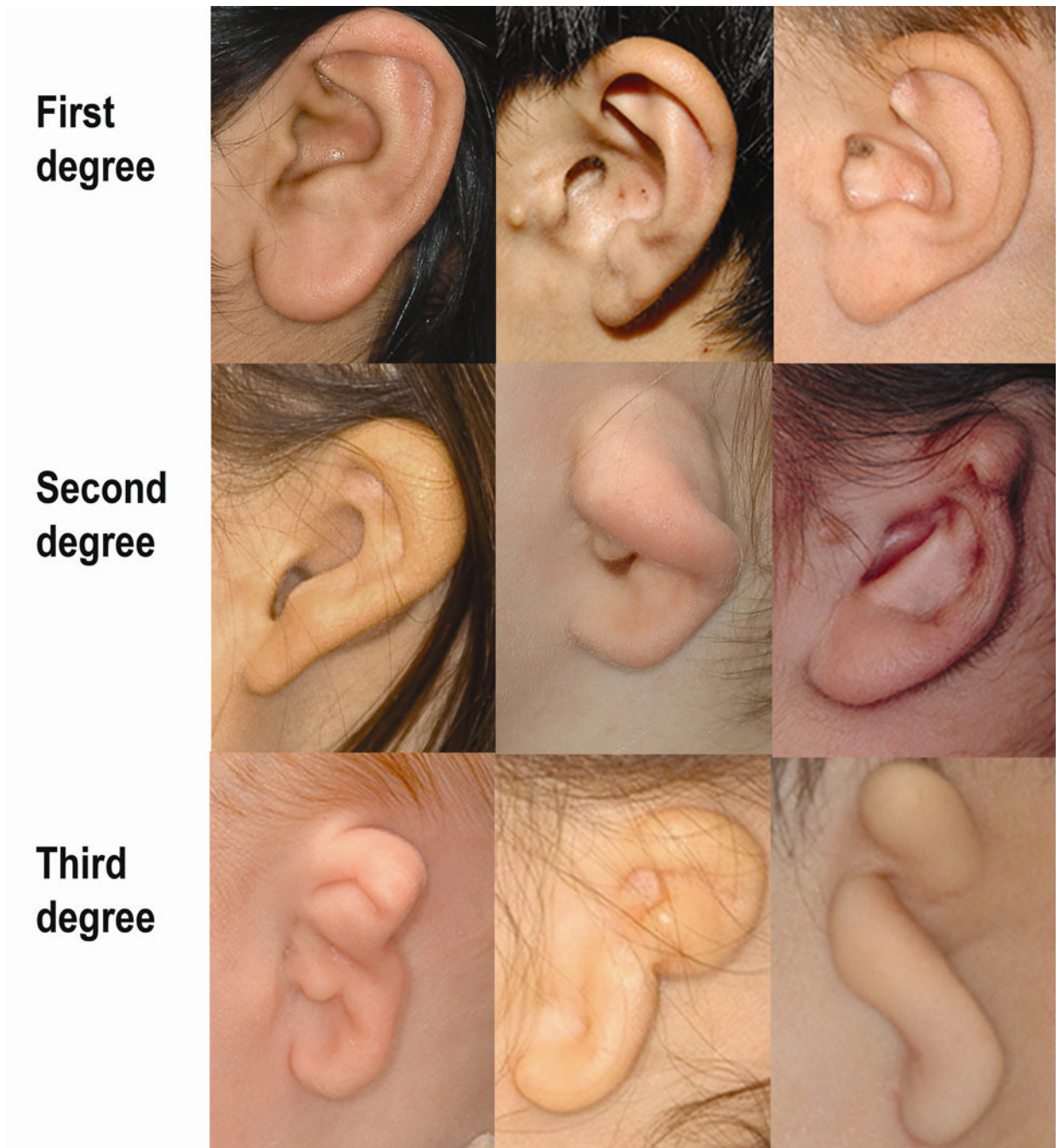
Figure 1.
Examples of ear photos for which all 6 raters agreed on the classification score. Top row: first degree microtia; middle row: second degree microtia and bottom row: third degree microtia.
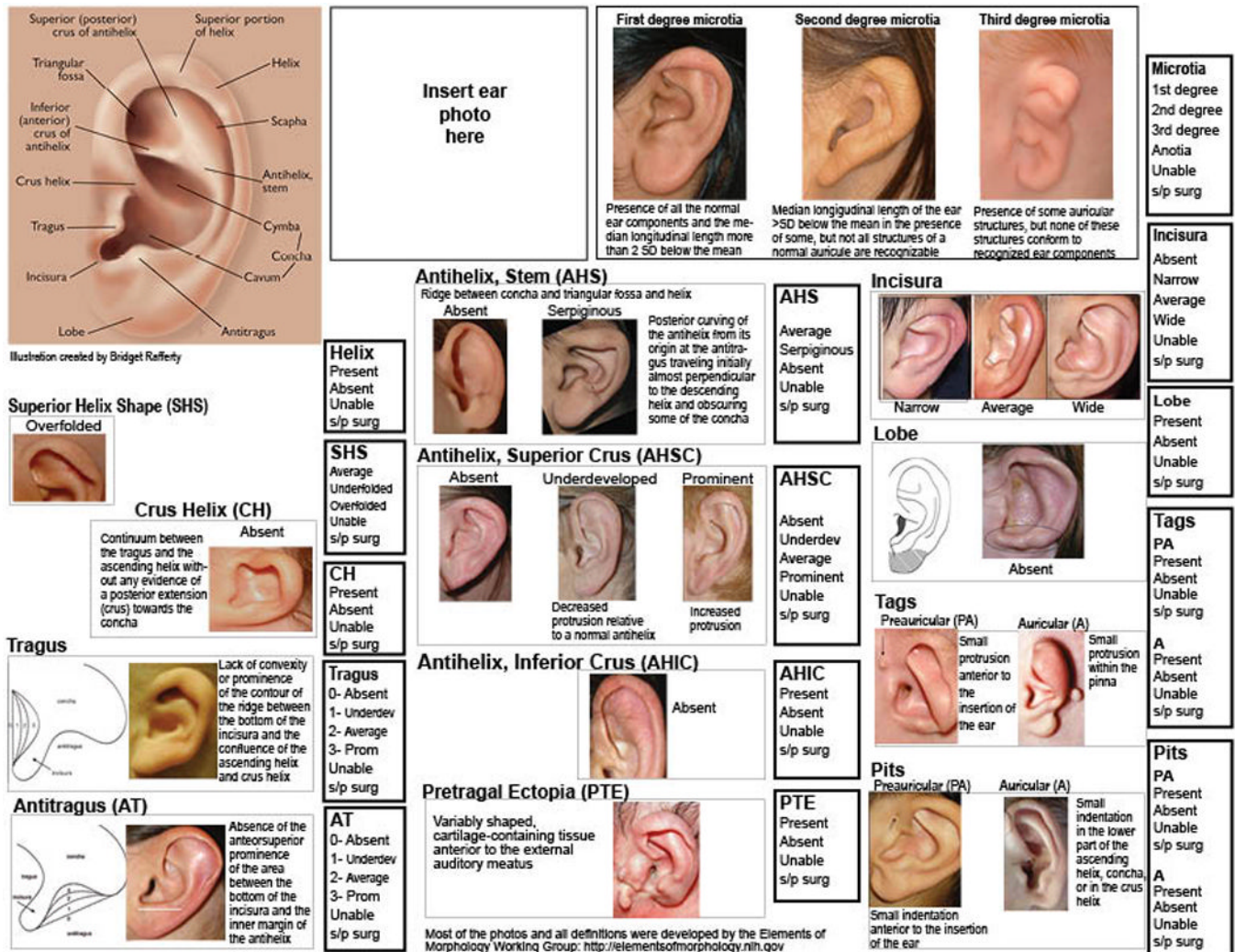
Figure 2.
PAT-Microtia: Phenotypic tool for evaluation of the morphology of the external ear observed in individuals with microtia. This tool incorporates written descriptions and illustrations based on the Elements of Morphology: Standard Terminology (EMST09), along with a standardized set of scoring options for each feature. Categorical scores are illustrated in the framed boxes.

Table I

Characteristics of the raters

	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6
Specialty	Plast	Ped	Oto	Ped	Plast	Oto
Years since fellowship	5–10	5–10	1–4	1–4	1–4	>20
Experience evaluating microtia	5–10	5–10	1–4	1–4	1–4	>20
Total time spent for rating (hours)	2	2.5	2.75	Unk	Unk	Unk

CF-Plast: Craniofacial Plastic Surgeon; Ped: Pediatrician; Oto: Otolaryngologist

Table II

General characteristics of the data, including the number of each ear feature rated, reasons indicated by raters for not rating, and number of missing data for each ear feature.

	Rated (N=600)	Unable to Rate due to[a]:			Missing[b]
		Absence of feature	Photo	Non-specified	
	n (%)	n (%)	n (%)	n (%)	n (%)
Helix					
Present/absent	563 (93.8)	0 (0.0)	9 (1.5)	3 (0.5)	25 (4.2)
Overall shape[b] (folded, under-, partially-, over-folded)	133 (22.2)	156 (26.0)	2 (0.3)	2 (0.3)	307 (51.2)
Shape of superior helix[b]	184 (30.7)	107 (17.8)	6 (1.0)	0 (0.0)	303 (50.5)
Antihelix, Stem					
Present/absent	567 (94.5)	0 (0.0)	13 (2.2)	6 (1.0)	14 (2.3)
Length	264 (44.0)	327 (54.5)	6 (1.0)	1 (0.2)	2 (0.3)
Prominence	251 (41.8)	326 (54.3)	20 (3.3)	1 (0.2)	2 (0.3)
Serpiginous	261 (43.5)	326 (54.3)	8 (1.3)	1 (0.2)	4 (0.7)
Antihelix, Inferior crus					
Present/absent	560 (93.3)	0 (0.0)	14 (2.3)	10 (1.7)	16 (2.7)
Length	235 (39.2)	348 (58.0)	11 (1.8)	3 (0.5)	3 (0.5)
Width	235 (39.2)	348 (58.0)	4 (0.7)	10 (1.7)	3 (0.5)
Prominence	229 (38.2)	348 (58.0)	17 (2.8)	3 (0.5)	3 (0.5)
Antihelix, Superior crus					
Present/absent	547 (91.2)	0 (0.0)	23 (3.8)	14 (2.3)	16 (2.7)
Width	177 (29.5)	396 (66.0)	6 (1.0)	19 (3.2)	2 (0.3)
Prominence	174 (29.0)	396 (66.0)	21 (3.5)	6 (1.0)	3 (0.5)
Antitragus					
Present/absent	558 (93.0)	0 (0.0)	8 (1.3)	3 (0.5)	31 (5.2)
Size	282 (47.0)	310 (51.7)	6 (1.0)	1 (0.2)	1 (0.2)
Crus helix					
Present/absent	572 (95.3)	0 (0.0)	6 (1.0)	5 (0.8)	17 (2.8)
Length	253 (42.2)	336 (56.0)	11 (1.8)	0 (0.0)	0 (0.0)

| | Rated (N=600) | Unable to Rate due to[a]: | | | Missing[b] |
		Absence of feature	Photo	Non-specified	
	n (%)	n (%)	n (%)	n (%)	n (%)
Prominence	255 (42.5)	336 (56.0)	9 (1.5)	0 (0.0)	0 (0.0)
Extension	101 (16.8)	342 (57.0)	0 (0.0)	0 (0.0)	157 (26.2)
Incisura					
Present/absent	570 (95.0)	0 (0.0)	11 (1.8)	3 (0.5)	16 (2.7)
Length	251 (41.8)	338 (56.3)	9 (1.5)	1 (0.2)	1 (0.2)
Width	251 (41.8)	338 (56.3)	11 (1.8)	0 (0.0)	0 (0.0)
Shape	224 (37.3)	337 (56.2)	17 (2.8)	0 (0.0)	22 (3.7)
Lobe (present/absent)	572 (95.3)	0 (0.0)	12 (2.0)	1 (0.2)	15 (2.5)
Tag (present/absent)	566 (94.3)	0 (0.0)	15 (2.5)	1 (0.2)	18 (3.0)
Pit (present/absent)	504 (84.0)	0 (0.0)	36 (6.0)	1 (0.2)	59 (9.8)
Tragus					
Present/absent	572 (95.3)	0 (0.0)	10 (1.7)	4 (0.7)	14 (2.3)
Size	299 (49.8)	290 (48.3)	6 (1.0)	0 (0.0)	5 (0.8)
Pretragal ectopia (present/absent)	572 (95.3)	0 (0.0)	2 (0.3)	4 (0.7)	22 (3.7)
Concha (normal/extra-fold)[b]	155 (25.8)	0 (0.0)	58 (9.7)	27 (4.5)	360 (60.0)
Marx classification	589 (98.2)	0 (0.0)	0 (0.0)	0 (0.0)	11 (1.8)

[a] According to the raters

[b] 300 of these missing ratings are due to absence of this category in the Batch 1 rating tool.

Table III

Prevalence of specific ratings for each component of 100 ear photographs independently rated by six raters (for a total of 600 possible ratings for each component).

Ear Components	N = 600* n (%)	Ear Components	N = 600* n (%)	Ear Components	N = 600* n (%)
Helix		*Antihelix superior crus*		*Lobe*	
Absent	240 (42.6)	Absent	359 (65.6)	Absent	71 (12.4)
Present	323 (57.4)	Present	188 (34.4)	Small	258 (45.1)
Shape of Helix		*Antihelix superior crus width*		Average	205 (35.8)
Folded	66 (49.6)	Narrow	3 (1.7)	Large	38 (6.6)
Underfolded	13 (9.8)	Average	64 (36.2)	*Tragus*	
Partially folded	16 (12.0)	Broad	110 (62.1)	Absent	266 (46.5)
Overfolded	38 (28.6)	*Antihelix superior crus prominence*		Present	306 (53.5)
Shape of Superior Helix		Underdeveloped	99 (56.9)	*Tragus size*	
Folded	108 (58.7)	Average	66 (37.9)	Underdeveloped	142 (47.5)
Underfolded	21 (11.4)	Prominent	9 (5.2)	Average	146 (48.8)
Partially folded	13 (7.1)	*Crus helix*		Prominent	9 (3.0)
Overfolded	42 (22.8)	Absent	315 (55.1)	Duplicated	2 (0.7)
Antihelix stem		Present	257 (44.9)	*Pretragal ectopia*	
Absent	301 (53.1)	*Crus helix length*		Absent	547 (95.6)
Present	266 (46.9)	Short	106 (41.9)	Present	25 (4.4)
Antihelix stem length		Average	116 (45.8)	*Concha*	
Short	104 (39.4)	Long	31 (12.3)	Normal	142 (91.6)
Average	134 (50.8)	*Crus helix prominence*		Extra folds	13 (8.4)
Long	26 (9.8)	Underdeveloped	73 (28.6)	*Antitragus*	
Antihelix stem prominence		Average	153 (60.0)	Absent	275 (49.3)
Underdeveloped	53 (21.1)	Prominent	29 (11.4)	Present	283 (50.7)
Average	152 (60.6)	*Crus helix extension*		*Antitragus size*	
Prominent	46 (18.3)	No	74 (73.3)	Absent	1 (0.4)
Antihelix stem serpigenous		Yes	27 (26.7)	Underdeveloped	92 (32.6)
No	186 (71.3)	*Incisura*		Average	119 (42.2)
Yes	75 (28.7)	Absent	312 (54.7)	Prominent	70 (24.8)

Ear Components	N = 600* n (%)	Ear Components	N = 600* n (%)	Ear Components	N = 600* n (%)	Ear Components	N = 600* n (%)
Antihelix inferior crus		Present	258 (45.3)	Present		*Tags*	
Absent	318 (56.8)	*Incisura length*		*Marx Classification*		Absent	472 (83.4)
Present	242 (43.2)	Short	89 (35.5)	Normal ear	108 (18.3)	Present	94 (16.6)
Antihelix inferior crus length		Average	138 (55.0)	Grade I	109 (18.5)	*Pits*	
Short	45 (19.1)	Long	24 (9.6)	Grade II	89 (15.1)	Present	65 (12.9)
Average	165 (70.2)	*Incisura width*		Grade III	244 (41.4)	Absent	439 (87.1)
Long	25 (10.6)	Wide	72 (28.7)	Grade IV	39 (6.6)		
Antihelix inferior crus width		Average	137 (54.6)				
Narrow	102 (43.4)	Narrow	35 (13.9)				
Average	122 (51.9)	Slit	7 (2.8)				
Broad	11 (4.7)	*Incisura shape*					
Antihelix inferior crus prominence		U	165 (73.7)				
Underdeveloped	45 (19.7)	V	59 (26.3)				
Average	151 (65.9)						
Prominent	33 (14.4)						

*
Totals do not always add to 600 because of missing values; the reasons for missing data is described in table 2.

Table IV

Interrater reliability overall and by medical specialty of **<u>each abnormality</u>** of the ear

Ear region	All	Plastic Surgeon	Otolaryngologist	Pediatrician
Specific ear component	ICC (95% CI)	ICC (95% CI)	ICC (95% CI)	ICC (95% CI)
Helix				
Present/absent	0.77 (0.72, 0.83)	0.81 (0.75, 0.88)	0.72 (0.62, 0.82)	0.79 (0.71, 0.87)
Overall shape (folded, under-, partially-, over-folded)	0.32 (0.13, 0.52)	0.35 (0.00, 0.76)	0.46 (0.12, 0.81)	0.00 (0.00, 0.50)
Shape of superior helix	0.63 (0.50, 0.77)	0.69 (0.49, 0.88)	0.60 (0.36, 0.83)	0.49 (0.19, 0.78)
Antihelix, Stem				
Present/absent	0.91 (0.88, 0.94)	0.87 (0.83, 0.92)	0.91 (0.88, 0.95)	0.93 (0.90, 0.96)
Length	0.32 (0.19, 0.46)	0.26 (0.00, 0.55)	0.33 (0.05, 0.60)	0.62 (0.43, 0.81)
Prominence	0.28 (0.15, 0.42)	0.48 (0.25, 0.72)	0.46 (0.20, 0.73)	0.23 (0.00, 0.53)
Serpiginous	0.37 (0.23, 0.51)	0.00 (0.00, 0.32)	0.54 (0.33, 0.75)	0.59 (0.38, 0.80)
Antihelix, Inferior crus				
Present/absent	0.93 (0.91, 0.95)	0.87 (0.82, 0.92)	0.96 (0.94, 0.97)	0.95 (0.93, 0.97)
Length	0.30 (0.15, 0.44)	0.35 (0.07, 0.64)	0.50 (0.25, 0.74)	0.21 (0.00, 0.52)
Width	0.04 (0.00, 0.14)	0.23 (0.00, 0.54)	0.22 (0.00, 0.53)	0.00 (0.00, 0.33)
Prominence	0.28 (0.13, 0.42)	0.53 (0.31, 0.76)	0.15 (0.00, 0.50)	0.01 (0.00, 0.34)
Antihelix, Superior crus				
Present/absent	0.81 (0.75, 0.86)	0.73 (0.63, 0.82)	0.83 (0.77, 0.90)	0.94 (0.92, 0.97)
Width	0.27 (0.11, 0.43)	0.33 (0.00, 0.69)	0.54 (0.27, 0.80)	0.00 (0.00, 0.43)
Prominence	0.48 (0.32, 0.65)	0.30 (0.00, 0.67)	0.27 (0.00, 0.64)	0.59 (0.33, 0.85)
Antitragus				
Present/absent	0.84 (0.80, 0.88)	0.85 (0.80, 0.91)	0.86 (0.81, 0.91)	0.90 (0.86, 0.94)
Size	0.53 (0.41, 0.66)	0.53 (0.31, 0.74)	0.62 (0.45, 0.79)	0.56 (0.34, 0.78)
Crus helix				
Present/absent	0.87 (0.83, 0.91)	0.94 (0.91, 0.96)	0.81 (0.74, 0.88)	0.85 (0.80, 0.91)
Length	0.36 (0.21, 0.50)	0.31 (0.02, 0.60)	0.60 (0.40, 0.79)	0.57 (0.36, 0.79)
Prominence	0.37 (0.23, 0.51)	0.31 (0.02, 0.60)	0.53 (0.31, 0.76)	0.30 (0.01, 0.60)
Extension	0.17 (0.00, 0.37)	0.00 (0.00, 0.51)	0.08 (0.00, 0.74)	0.17 (0.00, 0.66)
Incisura				
Present/absent	0.80 (0.75, 0.85)	0.92 (0.88, 0.95)	0.66 (0.55, 0.77)	0.88 (0.83, 0.93)
Length	0.23 (0.10, 0.37)	0.44 (0.20, 0.69)	0.00 (0.00, 0.40)	0.69 (0.52, 0.86)
Width	0.54 (0.41, 0.67)	0.80 (0.69, 0.91)	0.23 (0.00, 0.56)	0.72 (0.56, 0.88)
Shape	0.14 (0.01, 0.27)	0.03 (0.00, 0.37)	0.00 (0.00, 0.66)	0.28 (0.00, 0.60)
Lobe (present/absent)	0.50 (0.41, 0.59)	0.52 (0.37, 0.66)	0.55 (0.41, 0.70)	0.55 (0.40, 0.69)
Tag (present/absent)	0.76 (0.70, 0.82)	0.70 (0.60, 0.80)	0.85 (0.79, 0.91)	0.84 (0.77, 0.90)
Pit (present/absent)	0.45 (0.36, 0.55)	0.26 (0.06, 0.47)	0.29 (0.08, 0.50)	0.68 (0.55, 0.81)
Tragus				
Present/absent	0.84 (0.79, 0.88)	0.90 (0.86, 0.94)	0.87 (0.82, 0.92)	0.82 (0.75, 0.89)
Size	0.30 (0.17, 0.43)	0.07 (0.00, 0.36)	0.31 (0.05, 0.57)	0.58 (0.38, 0.78)
Pretragal ectopia (present/absent)	0.38 (0.28, 0.47)	0.65 (0.53, 0.76)	0.25 (0.06, 0.44)	0.35 (0.15, 0.54)

Ear region	All	Plastic Surgeon	Otolaryngologist	Pediatrician
Specific ear component	**ICC (95% CI)**	**ICC (95% CI)**	**ICC (95% CI)**	**ICC (95% CI)**
Concha (normal/extra-fold)	0.38 (0.20, 0.56)	0.51 (0.19, 0.82)	0.26 (0.00, 0.63)	0.53 (0.24, 0.81)
Marx classification	0.88 (0.85, 0.91)	0.93 (0.90, 0.95)	0.87 (0.82, 0.92)	0.85 (0.80, 0.91)