The Application of Clinical Genetics

Dovepress

open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Comparison of the Accuracy in Provisional Diagnosis of 22q11.2 Deletion and Williams Syndromes by Facial Photos in Thai Population Between De-Identified Facial Program and Clinicians

Nop Khongthon[1], Midi Theeraviwatwong[1], Khunton Wichajarn[2], Kitiwan Rojnueangnit[3]

[1]Medical Students, Faculty of Medicine, Thammasat University, Pathumthani, Thailand; [2]Department of Pediatrics, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand; [3]Department of Pediatrics, Faculty of Medicine, Thammasat University, Pathumthani, Thailand

Correspondence: Kitiwan Rojnueangnit, Division of Genetics, Department of Pediatrics, Faculty of Medicine, Thammasat University, 99/209 Moo 18 Phahonyothin Road, Klong-Luang, Pathumthani, 12120, Thailand, Tel +66-2-926-9514, Email rkitiwan@tu.ac.th

Introduction: There are more than 6000 genetic syndromes, therefore the recognition of facial patterns may present a challenge for clinicians. The 22q11.2 deletion syndrome (22q11.2 DS) and Williams syndrome (WS) are two different genetic syndromes but share some common phenotypic traits and subtle facial dysmorphisms. Therefore, any tool that would help clinicians recognize genetic syndromes would likely result in a more accurate diagnosis.

Methods: The syndrome identification accuracy was compared between 2 different facial analysis algorithms (DeepGestalt and GestaltMatcher) of the Face2Gene (F2G) tool and a group of 9 clinicians with different levels of expertise before and after using F2G for a cohort of 64 Thai participants' frontal facial photos divided into 3 groups of 22q11.2 DS, WS and unaffected controls.

Results: The higher accuracy from the DeepGestalt algorithm than from clinicians was demonstrated, especially when comparing between the two syndromes. The accuracy was highest when clinicians use the tool combined with their own decision-making process. The tool's second algorithm, GestaltMatcher revealed clear separation among these three groups of photos.

Discussion: The result of F2G outperforming clinicians was not surprising. However, the highest increase in accuracy was with nondysmorphology clinicians using F2G.

Conclusion: Face2Gene would be a useful tool to help clinicians in facial recognition of genetic syndromes, before ordering specific tests to confirm the definite diagnosis.

Keywords: 22q11.2 deletion syndrome, Williams syndrome, accuracy, GestaltMatcher, Face2Gene

Introduction

22q11.2 deletion syndrome (22q11.2 DS) is actually a common genetic condition with a prevalence of 1:2100–6000.[1–4] A small deletion of the long arm of chromosome 22 (locus q11.2, length 1.5–3 Mbp) interferes with the embryogenesis of the 3rd and 4th pharyngeal arches, which, consequently, can affect the formation of the conotruncal heart, thymus, and parathyroid. More than 180 associated phenotypes have been described,[5] with the common traits being delayed milestones, intellectual disability, behavioral issues, autism, as well as subtle facial features, such as auricular abnormalities, bulbous nose, hooded eyelids, widely spaced eyes, cleft palate, and asymmetric crying facies.[1,6]

Williams Syndrome (WS) is also a rare disease of genetic origin with a prevalence of 1:7500.[7] It is caused by a small deletion of the long arm of chromosome 7 (locus q11.23) or at the Williams-Beuren syndrome critical region (WBSCR) (length of 1.5–1.8 Mbp). The phenotypes can include congenital and acquired supravalvular aortic stenosis and/or peripheral pulmonary artery stenosis in the newborn period,[8] hypertension, hypercalcemia, delayed milestones, intellectual disability,

Received: 6 March 2024
Accepted: 31 May 2024
Published: 4 July 2024

overfriendly personality and facial dysmorphisms, eg, broad forehead, bitemporal narrowing, periorbital fullness, stellate or lacy iris pattern, strabismus, short nose, midface hypoplasia, lower full cheeks, long philtrum, full vermilion lips, macrostomia, small, and widely spaced teeth.[9,10] Although new technologies are available in genetic diagnosis, facial dysmorphism is currently still an important tool in genetic evaluation.[11] These two syndromes have some common sharing systems involvement, include cardiac defects, delayed milestones and facial dysmorphism, such as eyelid fullness or broad nose, which are subtle, especially in younger children. The definite diagnosis can be made currently by genetic testing but in the absence facial dysmorphism, clinicians may not order any genetic tests. On the other hand, if clinicians recognize the dysmorphisms and suspect a specific syndrome, the precise test could be performed, reducing costs and waiting time. Such is the example of performing a fluorescent in situ hybridization (FISH) test to detect the specific deletion when WS or 22q11.2 DS are suspected. However, if clinicians are aware of dysmorphism, but are unable to recognize specific syndromes; more than one test or chromosomal microarray will probably be performed, leading to at least double the cost (or as high as 4–8 times more) than one specific FISH test (based on the cost of the test in Thailand and the rest of the world).[12,13] These genetic tests can be ordered by any clinician or pediatrician; the referral to clinical geneticists is another option but is available in only some tertiary hospitals. Therefore, if there are tools to help clinicians, not specialized in dysmorphology, better identify suspected syndromes, this would increase the precision for diagnosis, and reduce the need for unnecessary investigation costs.

The Face2Gene (FDNA Inc USA) platform uses artificial intelligence algorithms that help clinicians in genetic syndromic recognition. These algorithms use 2-dimension images of the person's face and convert the image by using deep convolutional neutral networks (DCNNs) to a mathematical facial descriptor, which de-identifies the image. In one of these algorithms named DeepGestalt, this facial descriptor is compared to syndrome gestalts to quantify similarity, resulting in a prioritized list of the top 30 syndromes with similar morphology.[14,15] However, this method can recognize only 360 syndromes (vs.22.1.0) that the technology was already trained to identify, with the limitation that it is unable to recognize syndromes that it has not been trained in, or unable to recognize novel syndromes.[16] Although DeepGestalt is not trained to identify photos of unaffected faces, facial dysmorphism has been assessed and the level of dysmorphic features are correlated by dysmorphic (D-score) for pediatrician view. Recently, the GestaltMatcher algorithm has been added to the platform helping to match similar photos together either for known or unknown diagnosis. The process is started by encoding a patient photo into a 320-dimensional vector, called Facial Phenotype Descriptor (FPD), which can be represented as a coordinate in the Clinical Face Phenotype Space (CFPS). Distances between different FPDs in the CFPS define syndromic similarity – smaller distances mean higher similarity. FPD also converts to 2-dimensions, which can visualize them in a scatterplot, called t-distributed stochastic neighbor embedding (t-SNE) visualizations.[16] In a pairwise comparison matrix, we compared our cases (X and Y axes) images with themselves and with the GestaltMatcher gallery (4300 images), outputting similarity ranks. The dark green values (low ranks) in this matrix indicated our groups of patients were more similar among themselves than when compared to the gallery images.

There has been much research using Face2Gene (F2G) as a tool for providing information about genetic diagnosis based on facial gestalt. One such study demonstrated a compelling result distinguishing between Emanuel and Pallister-Killian syndromes, which are caused by small supernumerary marker chromosomes, and also excelled in differentiating between these two syndromes and photos of unaffected controls or controls of other syndromic facies.[17] Another study used F2G for diagnosing genetic diseases from two textbook photos, which revealed around a 42–54% detection rate if the cut-point was at the first and third rank, respectively.[18] 22q11.2 deletion[19] and Williams syndrome[20] were previously studied in diverse populations, nevertheless only 8–17% were Asian, and only 2 and 4 facial photos are of Thai people.

Facial features depend on ancestral background; many studies have been conducted in Caucasian populations, fewer studies were done in Asian people and only Down syndrome has been studied in Thai population.[21,22] Therefore, our two main objectives in this study included: whether these algorithms of F2G can be utilized for other genetic syndromes, ie, 22q11.2 DS and WS in Thai population; and the 2G can be used as a screening tool helping non-dysmorphology clinicians recognize any genetic syndromes.

Materials and Methods

All participants' frontal facial photos were taken by the researchers, one photo for each person. The diagnosis of 22q11.2 DS and WS were confirmed by FISH test. Unaffected children were selected from children without any underlying disease and screening for normal development in Well-Child Care Clinic.

Our study involved clinicians (divided into three subgroups: interns, pediatric residents, and pediatricians) who were not specialized in dysmorphology, to evaluate frontal facial photographs from three distinct clinical groups (which were also divided into three subgroups: patients with 22q11.2 DS, patients with WS, and unaffected controls). This study was done at Thammasat and Khon Kaen Universities during January–May 2021 and approved by the Human Research Ethics Committee of Thammasat University No 1 (Faculty of Medicine) (MTU-EC-SA-0-031/63), and the Khon Kaen University Ethics Committee for Human Research, Panel 1 (HE631606). This study complied with the declaration of Helsinki regarding research on humans. Written informed consent was obtained from each individual and/or their parents if they or their children would like to participate in the study by having their facial photos taken and publication of the images.

All frontal facial photos were uploaded to F2G CLINIC, for both non-Pediatrician and Pediatrician view to evaluate facial D-score, with the top-3, top-10, and top-30 ranks of syndromes-matches by the DeepGestalt algorithm being recorded. In addition, we utilized the F2G RESEARCH app to analyze the receiver operating characteristic (ROC) curve and multiclass matrix.

In parallel, all facial photos were presented to clinicians who needed to decide whether it was 22q11.2 DS, WS, or an unaffected child for each photo individually twice. These clinicians were blinded to the patients' chart or FISH test result. The first decision was made independent by themselves, then the second decision was made again after having the suggested genetic syndrome from the F2G CLINIC individuallyeither 22q11.2 DS or WS. We used the top-3 for the cut-off point, either 22q11.2 DS, WS, or not both 22q11.2 DS and WS, given this is predicted to be the highest accuracy, a hypothesis supported by the result of the previous study.[21]

We combined the clinicians' results in each subgroup; the answer for 22q11.2 DS, WS, or unaffected children would be marked if two or more answers were similar. The comparison was performed for sensitivity, specificity and accuracy between F2G and clinicians before and after use. The methodology of our study is shown in Figure 1.

These photos were also analyzed by the GestaltMatcher algorithm for each subgroup of 22q11.2 DS, WS and unaffected controls. We performed a pairwise rank matrix analysis and t-SNE.

Results

Sixty-four participants were enrolled, divided into three subgroups: 16 patients with 22q11.2 DS, 16 patients with WS, and 32 unaffected children. Nine clinicians were enrolled to our study, also divided into three groups: 3 interns, 3 pediatric residents, and 3 pediatricians. All demographic data of participants are demonstrated in Supplemental Table 1.

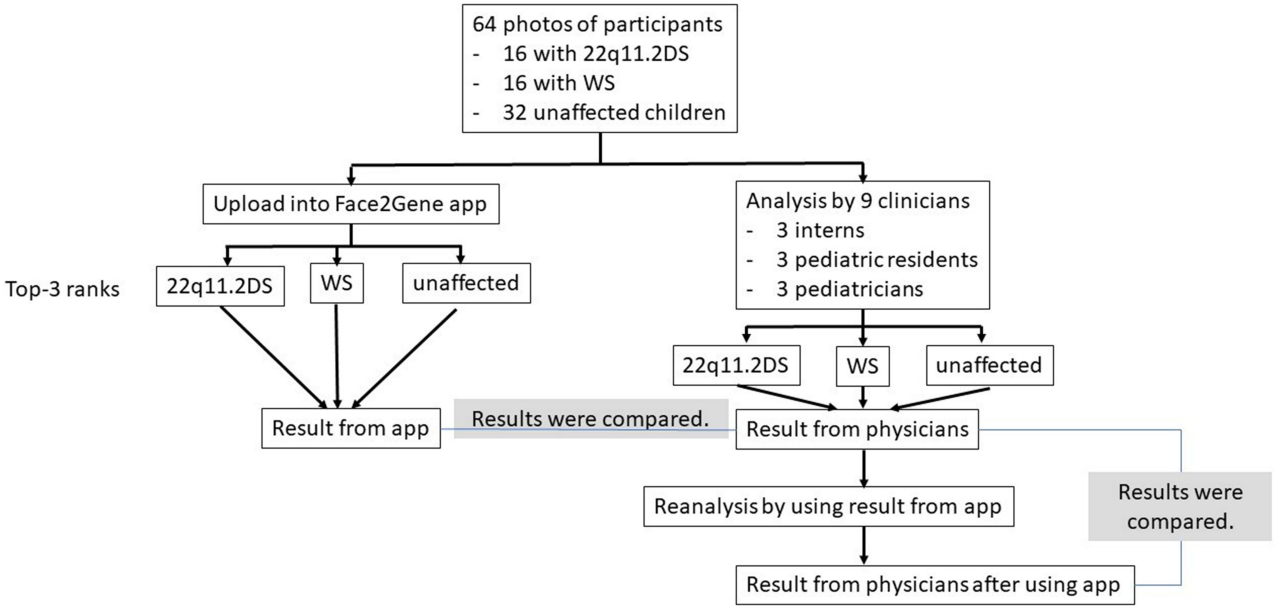

Figure 1 Methodology of our study.

Our study confirmed the top-three ranks were the most appropriate cut-off for syndrome recognition by the highest accuracy with the area under a ROC curve of 0.8867, when compared with the top-10, and the top-30 for cut-off point with the area under ROC curve of 0.8364, and 0.6852, respectively (Supplementary Figure 1). The result of each photo analyzed by F2G's DeepGestalt is shown in Supplemental Table 2. Subgroup analysis for the unaffected group revealed more than half (56%) of the participants had high D-scores, which mean high possible for having dysmorphism, but only 8 (25%) participants had low D-scores, or low possible having facial dysmorphism (Supplemental Table 3).

The comparison between F2G's DeepGestalt and clinicians' results for 22q11.2 DS group demonstrated that the accuracy of DeepGestalt (97%) was higher than clinicians for the syndromic recognition (75%), when compared with WS group; sensitivity and specificity of the F2G were 94 and 100%, respectively, which was also higher than clinicians (69 in sensitivity and 75% in specificity). However, clinicians had higher accuracy (83%) and specificity (91%) for recognition of 22q11.2 than F2G (79 and 72%, respectively) compared with the unaffected controls. For WS group, F2G demonstrated 100% accuracy, sensitivity and specificity in recognition when compared with 22q11.2 DS group, and 98% accuracy when compared with the unaffected controls. Clinicians were lower in all results, accuracy (75–88%), sensitivity (81%) and specificity (69–91%) (Table 1).

Regarding the correlation in each group of clinicians before using the app, pediatric residents seemed to have the highest accuracy (84%) for recognition of both 22q11.2 DS and WS when comparing with these syndromes together but of no statistical significance. There was also no difference in the accuracy among each group of clinicians when compared with the unaffected controls. When clinicians utilized the results from F2G, all results of suspected syndromes increased in accuracy, sensitivity and specificity. Interestingly, the intern group received the highest increase in score in terms of accuracy, followed by pediatric resident, and pediatrician groups, respectively. The overall accuracy in recognition for both syndromes before and after using the app is presented in Table 2.

The multiclass comparison in F2G RESEARCH app demonstrates 95% sensitivity for WS, which is correlated with WS group, and 100% sensitivity for unaffected group, while there was only 61% sensitivity for 22q11.2 DS. Conversely, 38% of 22q11.2 DS was recognized as unaffected, while only 5% of WS was recognized as unaffected (Table 3).

The binary comparison results demonstrate strong discrimination among those with 22q11.2 DS and WS; WS and unaffected group, confirmed by an AUC of 1, whereas 22q11.2 DS and unaffected was 0.93, as shown in Figure 2.

Table 1 Comparison of the Sensitivity, Specificity and Accuracy in Recognition of 22q11.2 Deletion Syndrome and Williams Syndrome with Other Groups Between Face2Gene and Clinicians

	22q11.2 DS Compared with	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**Accuracy (95% CI)**
Face2Gene	WS	94 (70–100)	100	97 (84–100)
	Unaffected		72 (53–86)	79 (65–90)
Clinicians	WS	69 (41–88)	75 (47–92)	75 (56–88)
	Unaffected		91 (74–98)	83 (69–92)
	WS compared with	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**Accuracy (95% CI)**
Face2Gene	22q11.2 DS	100	100	100
	Unaffected		97 (84–100)	98 (90–100)
Clinicians	22q11.2 DS	81 (54–95)	69 (41–88)	75 (56–88)
	Unaffected		91 (74–98)	88 (74–95)

Abbreviations: 22q11.2DS, 22q11.2 deletion syndrome; CI, confidence interval; WS, Williams syndrome.

Table 2 Comparison of the Sensitivity, Specificity and Accuracy in Recognition of 22q11.2 Deletion Syndrome and Williams Syndrome in Each Group of Clinicians Before and After Utilizing Recognition Pattern Results of Face2Gene Application

		Sensitivity (95% CI)		Specificity (95% CI)		Accuracy (95% CI)	
22q11.2 DS Compared with		**WS**	**Unaffected**	**WS**	**Unaffected**	**WS**	**Unaffected**
Pediatricians	Before	56 (30–80)		81 (54–96)	91 (75–98)	69 (50–83)	79 (65–90)
	After	63 (35–85)		88 (62–98)	91 (75–98)	75 (57–89)	81 (67–91)
Pediatric residents	Before	94 (70–100)		75 (48–93)	72 (53–86)	84 (67–95)	79 (65–90)
	After	100 (79–100)		94 (70–100)	78 (60–91)	97 (84–100)	85 (72–94)
Interns	Before	63 (35–85)		69 (41–89)	94 (79–99)	66 (47–81)	83 (70–93)
	After	75 (48–98)		94 (70–100)	97 (84–100)	84 (67 −95)	90 (77–97)
Williams syndrome compared with		**22q11.2 DS**	**Unaffected**	**22q11.2 DS**	**Unaffected**	**22q11.2 DS**	**Unaffected**
Pediatricians	Before	81 (54–95)		56 (29–80)	91 (74–98)	69 (49–83)	88 (76–94)
	After	88 (61–98)		63 (35–84)	91 (74–98)	75 (56–88)	90 (77–96)
Pediatric residents	Before	75 (47–92)		94 (69–99)	72 (58–84)	84 (67–94)	73 (58–84)
	After	94 (69–99)		100 (0.1–30)	78 (60–90)	96 (83–99)	83 (69–92)
Interns	Before	69 (41–88)		63 (35–84)	94 (79–99)	66 (46–81)	85 (72–93)
	After	94 (69–99)		75 (47–92)	97 (83–99)	84 (67–94)	96 (85–99)

Abbreviations: 22q11.2DS, 22q11.2 deletion syndrome; CI, confidence interval; WS, Williams syndrome.

Table 3 The Multiclass Comparison Among 22q11.2 Deletion, Williams Syndrome and Unaffected Groups

Predicted / Actual	Unaffected	22q11.2 Deletion Syndrome	Williams Syndrome
Unaffected	1	0	0
22q11.2 deletion syndrome	0.38	0.61	0.01
Williams syndrome	0.05	0	0.95

From the mathematical facial descriptor illustrated as a composite image by F2G for each group, it revealed narrow palpebral fissures and subtle bulbous nose in 22q11.2 DS group when compared with the unaffected controls. However, WS group revealed more unique facial features including swollen eyelids, full and lower cheeks with macrostomia (Figure 3).

In GestaltMatcher algorithm, 63 out of 64 photos were able to generate these vectors. One photo failed, which might have happened due to poor image quality. The pairwise matrix of our 63 images were compared with the F2G gallery of 4300 images. The matrix's results revealed a well-defined cluster for each group of 22q11.2 DS, WS and unaffected controls; meaning each subgroup was more similar amongst themselves than to other photos in the gallery (Figure 4A).

The t-SNE visualizations, where a dimensionality reduction method was used to convert the 320-dimension FPDs to 2-dimensions to visualize them in a 2D scatterplot, revealed three separate clusters without overlapping, indicating these three subgroups are similar themselves but different from other groups (Figure 4B). The confidence ellipses depicted have a radius of one standard deviation from each cluster's centroid.

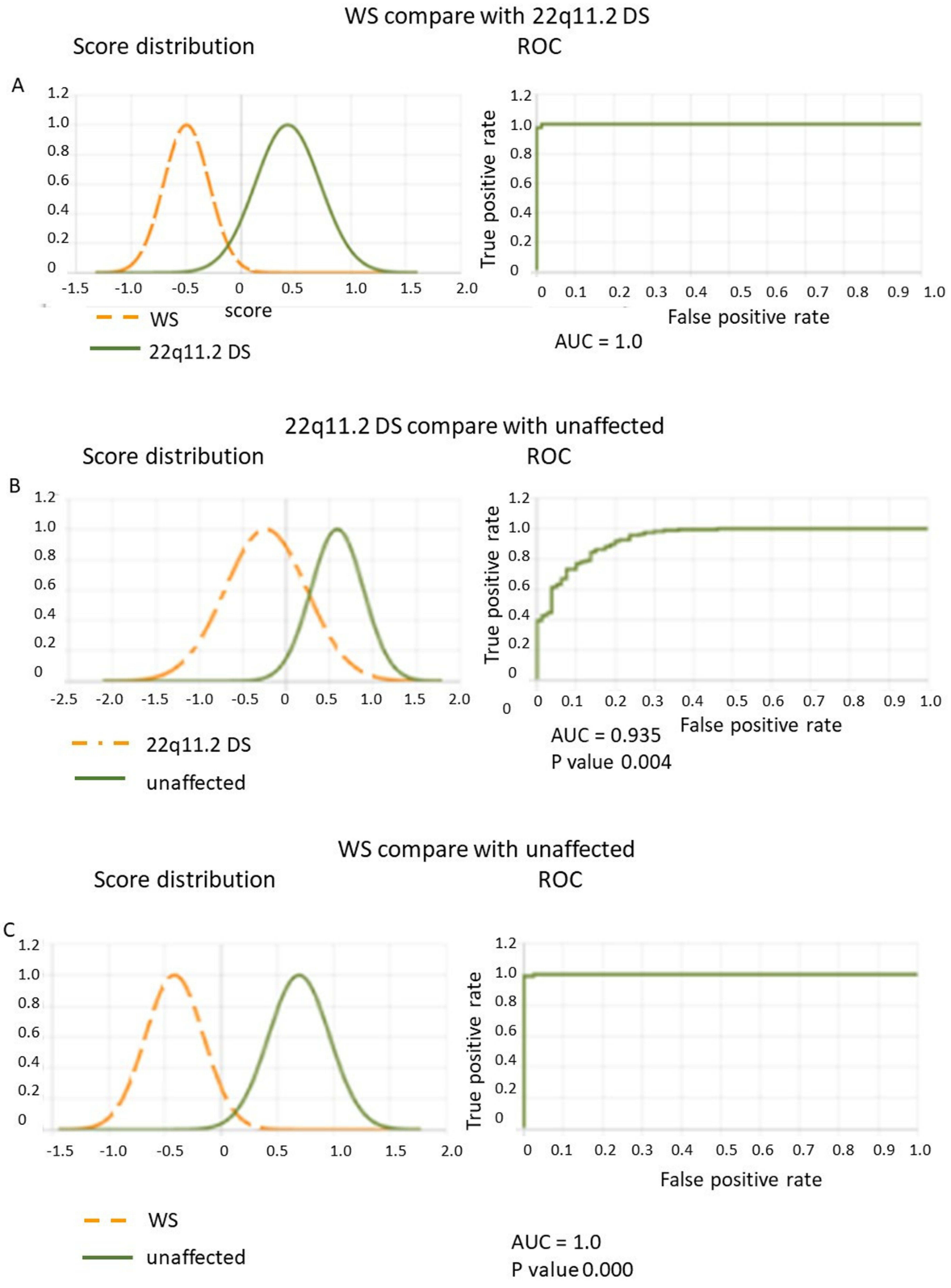

Figure 2 Binary comparison demonstrating (**A**) left: FDNA technology detection score curves of the WS (in dot) and 22q11.2 DS (in line). The x-axis represents the score and y-axis represents score distribution. The curves for each group are minimal overlap, confirmed by (**A**) right: area under curve (AUC) of 1.0; (**B**) score curves of the 22q11.2 DS (in dot) and unaffected (in line), the curves partially overlap, with the AUC of 0.93 and *P*-value of 0.004; (**C**) score curves of the WS (in dot) and unaffected (in line), the curves minimally overlap, with the AUC of 1.0.

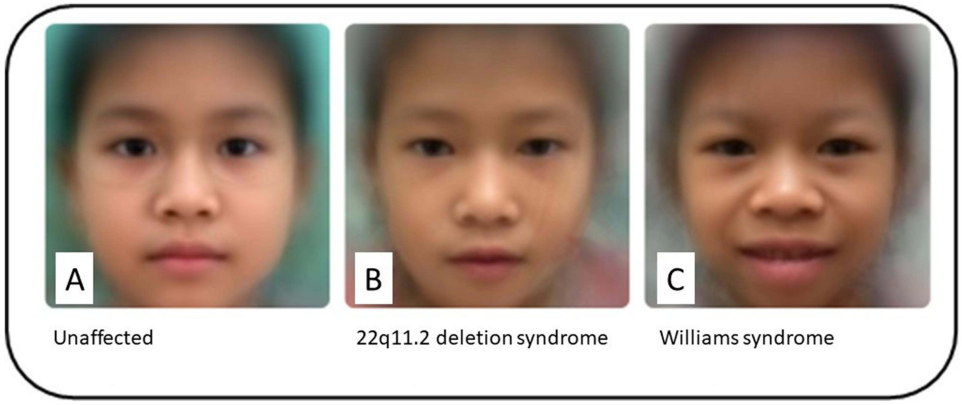

Figure 3 Artificial intelligence of facial descriptor illustrated by Face2Gene application of the group of (**A**) unaffected, (**B**) 22q11.2 deletion syndrome and (**C**) Williams syndrome.

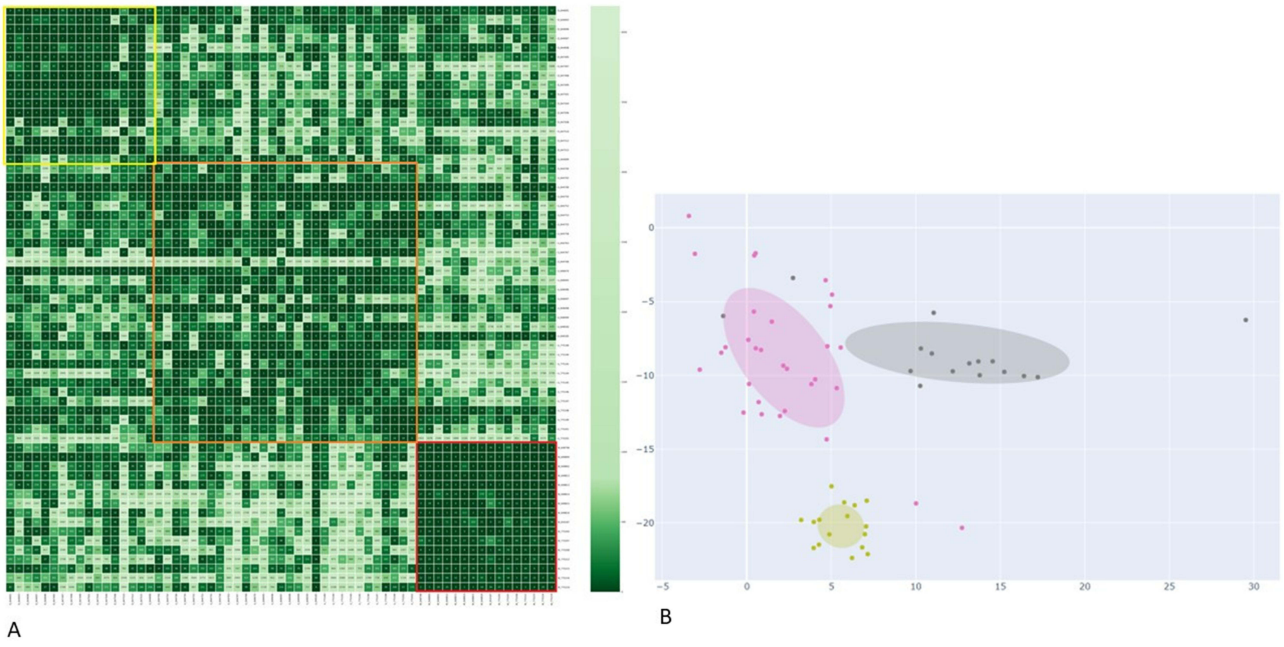

Figure 4 (**A**) Pairwise ranks of individuals in our cohort. The case identifier number (ID) is on both Y and X axes in the same order, therefore, zeros are on the diagonal, since it is the similarity rank achieved when each photo is compared to itself. Yellow square represents 22q11.2 deletion syndrome group, Orange square represents unaffected group and red square represents Williams syndrome group. The dark green values indicate our groups are more similar. (**B**) t-distributed stochastic neighbor embedding (t-SNE) visualizations reveals three distinct clusters in the plot without overlapping. Grey ellipse represents 22q11.2 deletion syndrome group, pink represents unaffected, while yellow represents Williams syndrome.

Discussion

F2G better identified both 22q11.2 DS and WS with higher accuracy than clinicians. It demonstrated excellent performance in suspected WS with 100% sensitivity, specificity and accuracy. Similarly, the performance in suspected 22q11.2 DS was also strong in accuracy (97%) when compared between these two syndromes. Not surprisingly, that F2G was more accurate in recognizing genetic syndromes compared with clinicians, given these two syndromes were not too rare and part of the F2G DeepGestalt training.[19,20] Clinicians in this study were interns, pediatric residents, and general pediatricians, who are not specialized in dysmorphology. We chose this group of clinicians as they are usually the primary doctors for these children.

This comparison between F2G and clinicians was unbalanced, given that the suggestive syndrome from F2G DeepGestalt was one in the 360 syndromes for which the algorithm was trained for. Nevertheless, the clinicians had

only three choices of 22q11.2 DS, WS, or unaffected for an answer. This also confirmed the superiority of artificial intelligence algorithms in recognizing facial features.

We were initially concerned that the facial features of 22q11.2 DS might overlap with unaffected Asian faces. This hypothesis was confirmed by the result of our unaffected controls were recognized as 22q11.2 DS in 34% (11/32) by the top-three ranks and high as 53% (17/32) by using top-10 ranks. However, it must be noted that DeepGestalt is not trained to identify photos of unaffected children. Additionally, the multiclass comparison also demonstrated that the facial features of 22q11.2 DS might be similar to the unaffected group by 38%, at the AUC of 0.935. This binary comparison indicated that although the app might be confused between 22q11.2 DS and unaffected faces, there was efficacy in separation between these two groups. When assessing the level of dysmorphology with Face2Gene by facial D-score for pediatrician view in unaffected group, 9/11 photos in the top-three ranks presented with higher D-score, while another 2 presented with in middle (unknown significant) score. Only 25% (8/32) photos of unaffected group presented with low D-score, which support the less likely possible for genetic syndrome. These findings support the hypothesis that subtle facial features of 22q11.2 DS may be confused with unaffected Asian faces. An inadequate number of Thai patients' facial photos with 22q11.2 DS in app training might be an issue.[19] However, F2G GestaltMatcher algorithm, although there is variation in terms of facial gestalt in each group, this algorithm can separate both 22q11.2 DS and unaffected group, without dependence on app training.

In contrast, WS facial features are unique and easier to separate from unaffected controls, as the result showed greater accuracy in recognition by both clinicians and F2G, which none of unaffected children was recognized as WS by the top-three ranks, in confusion matrix, binary comparison as well as in precise results in GestaltMatcher algorithm (dark green in pairwise and t-SNE visualizations). Therefore, our hypothesis about the facial similarity between WS and 22q11.2 DS was erroneous, given the result demonstrated distinguishing differences between these two syndromes.

Our result of an increase in syndromic genetic recognition after using F2G demonstrated the benefit of using this tool in a clinical setting to help clinicians, not trained in dysmorphology, recognize genetic syndromes. One limitation of this study was only a few clinicians participating, it was interesting to note that the intern subgroup, who had less clinical experience, had more accuracy in syndromic recognition after using F2G than both residents and pediatricians, who have more clinical experience. On the other hand, the outcome that pediatric residents seemed to perform the best in recognition among the three groups of clinicians was explained by residents having the most up-to-date knowledge about dysmorphic features due to their training period.

Conclusion

Therefore, our study supports the notion that F2G is a useful tool for genetic syndromic suggestion with both algorithms, and our study demonstrated the utilization of the tool in a clinical setting can help nondysmorphology clinicians with syndromic recognition. However, the final diagnosis always requires confirmation from genetic tests.

Acknowledgments

We thank patients and their families who contributed photographs, all physicians who participated in the study, Prof. Dr. Paskorn Sritipsukho for helping in statistical analysis, the TU Dome Med Star program for facilitating the research, Ms. Sam Ormond, a native English speaker for editing this manuscript, Ms. Nicole Fleischer and Ms. Carolina Alves from FDNA's research collaboration for helping in using the research part analysis of Face2Gene.

Author Contributions

Nop Khongthon: designed the study protocol, collected and analysed the data, wrote the draft and approved of the manuscript. Midi Theeraviwatwong: collected the data and approved of the manuscript. Khunton Wichajarn: collected the data, edited and approved of the manuscript. Kitiwan Rojnueangnit: designed the study protocol, analyzed data, wrote, and approved the manuscript. All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Disclosure

All authors report no conflicts of interest in this work.

References

1. McDonald-McGinn DM, Sullivan KE, Marino B, et al. 22q11.2 deletion syndrome. *Nat Rev Dis Primers*. 2015;1:15071. doi:10.1038/nrdp.2015.71
2. Botto LD, May K, Fernhoff PM, et al. A population-based study of the 22q11.2 deletion: phenotype, incidence, and contribution to major birth defects in the population. *Pediatrics*. 2003;112(1 Pt 1):101–107. doi:10.1542/peds.112.1.101
3. Oskarsdóttir S, Vujic M, Fasth A. Incidence and prevalence of the 22q11 deletion syndrome: a population-based study in Western Sweden. *Arch Dis Child*. 2004;89(2):148–151. doi:10.1136/adc.2003.026880
4. Blagojevic C, Heung T, Theriault M, et al. Estimate of the contemporary live-birth prevalence of recurrent 22q11.2 deletions: a cross-sectional analysis from population-based newborn screening. *CMAJ Open*. 2021;9(3):E802–E809. doi:10.9778/cmajo.20200294
5. Robin NH, Shprintzen RJ. Defining the clinical spectrum of deletion 22q11.2. *J Pediatr*. 2005;147(1):90–96. doi:10.1016/j.jpeds.2005.03.007
6. McDonald-McGinn DM, Emanuel BS, Zackai EH. 22q11.2 deletion syndrome. In: Adam MP, Ardinger HH, Pagon RA, et al. editors. *GeneReviews®*. Seattle: University of Washington, Seattle; 2015. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1523/. Accessed April 3, 2023.
7. Strømme P, Bjørnstad PG, Ramstad K. Prevalence estimation of Williams syndrome. *J Child Neurol*. 2002;17(4):269–271. doi:10.1177/088307380201700406
8. Pober BR, Johnson M, Urban Z. Mechanisms and treatment of cardiovascular disease in Williams-Beuren syndrome. *J Clin Invest*. 2008;118(5):1606–1615. doi:10.1172/JCI35309
9. Morris CA. Williams Syndrome. In: Adam MP, Everman DB, Mirzaa GM, et al. editors. *GeneReviews®*. Seattle: University of Washington, Seattle; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1249/. Accessed April 3, 2023.
10. Morris CA. Introduction: Williams syndrome. *Am J Med Genet C Semin Med Genet*. 2010;2010:154.
11. Hurst ACE, Robin NH. Dysmorphology in the era of genomic diagnosis. *J Pers Med*. 2020;10(1):18. doi:10.3390/jpm10010018
12. Regier DA, Friedman JM, Marra CA. Value for money? Array genomic hybridization for diagnostic testing for genetic causes of intellectual disability. *Am J Hum Genet*. 2010;86(5):765–772. doi:10.1016/j.ajhg.2010.03.009
13. Chromosome microarray. hfocus; 2022. Available from: https://www.hfocus.org/content/2016/04/11990. Accessed January 6, 2023.
14. Gurovich Y, Hanani Y, Bar O, et al. Identifying facial phenotypes of genetic disorders using deep learning. *Nat Med*. 2019;25(1):60–64. doi:10.1038/s41591-018-0279-0
15. INC, F. (2011–2022). FACE2GENE. Available from: https://www.face2gene.com/. Accessed January 6, 2023.
16. Hsieh TC, Bar-Haim A, Moosa S, et al. GestaltMatcher facilitates rare disease matching using facial phenotype descriptors. *Nat Genet*. 2022;54(3):349–357. doi:10.1038/s41588-021-01010-x
17. Liehr T, Acquarola N, Pyle K, et al. Next generation phenotyping in Emanuel and Pallister-Killian syndrome using computer-aided facial dysmorphology analysis of 2D photos. *Clin Genet*. 2018;93(2):378–381. doi:10.1111/cge.13087
18. Javitt MJ, Vanner EA, Grajewski AL, Chang TC. Evaluation of a computer-based facial dysmorphology analysis algorithm (Face2Gene) using standardized textbook photos. *Eye*. 2022;36(4):859–861. doi:10.1038/s41433-021-01563-5
19. Kruszka P, Addissie YA, McGinn DE, et al. 22q11.2 deletion syndrome in diverse populations. *Am J Med Genet A*. 2017;173(4):879–888. doi:10.1002/ajmg.a.38199
20. Kruszka P, Porras AR, de Souza DH, et al. Williams-Beuren syndrome in diverse populations. *Am J Med Genet A*. 2018;176(5):1128–1136. doi:10.1002/ajmg.a.38672
21. Vorravanpreecha N, Lertboonnum T, Rodjanadit R, Sriplienchan P, Rojnueangnit K. Studying Down syndrome recognition probabilities in Thai children with de-identified computer-aided facial analysis. *Am J Med Genet A*. 2018;176(9):1935–1940. doi:10.1002/ajmg.a.40483
22. Srisraluang W, Rojnueangnit K. Facial recognition accuracy in photographs of Thai neonates with Down syndrome among physicians and the Face2Gene application. *Am J Med Genet A*. 2021;185(12):3701–3705. doi:10.1002/ajmg.a.62432