Scientific Reports

https://doi.org/10.1038/s41598-025-30593-2

Article in Press

Characteristics of fetal facial expression changes using artificial intelligence: a pilot study

Received: 26 September 2025

Accepted: 26 November 2025

Published online: 28 November 2025

Cite this article as: Miyagi Y., Hata T. & Miyake T. Characteristics of fetal facial expression changes using artificial intelligence: a pilot study. *Sci Rep* (2025). https://doi.org/10.1038/s41598-025-30593-2

Yasunari Miyagi, Toshiyuki Hata & Takahito Miyake

REVISED Submission ID 7b3c9f08-a446-4e03-adc5-bd234b49e8b3

ARTICLE

Characteristics of fetal facial expression changes using artificial intelligence – A pilot study

Yasunari Miyagi [1,2*], Toshiyuki Hata [3], Takahito Miyake [1,3]

1. Department of Obstetrics and Gynecology, Miyake Ofuku Clinic, Okayama city, Japan
2. Medical Data Labo, Okayama city, Japan
3. Department of Obstetrics and Gynecology, Miyake Clinic, Okayama city, Japan

*Corresponding author:
Yasunari Miyagi, MD, PhD

1. Miyake Ofuku Clinic

393-1 Ofuku, Minami ward, Okayama city, Okayama prefecture 701-0204, Japan

Tel.: +81-86-281-2020

e-mail: ymiyagi@mac.com

2. Medical Data Labo

289-48 Yamasaki, Naka ward, Okayama, Okayama prefecture 703-8267, Japan

Abstract

We aimed to investigate the frequency, changes, and chaotic correlation dimensions of fetal facial expression videos using artificial intelligence (AI) and speculate the state of the fetal brain activity. We applied our original AI for classifying fetal facial expressions to 57,208 frames, total of 95.27 minutes, from 47singleton pregnancies at 28 to 37 weeks of gestation, obtained at Miyake Clinic between December 2023 and February 2024 at 0.1-second intervals. Time, transitions and correlation dimensions of the facial expressions were investigated. There was a significant difference between expressions. Neutral and mouthing showed significantly longer durations; 71.0, 9.4 – 174.8 (Mean, 5 – 95 %ile) and 53.3, 0.7 – 127.3 seconds for neutral and mouthing, respectively. The longest transitions were neutral to mouthing at 2,237.5 seconds. The median correlation dimensions for before, during, and after neutral and mouthing were 1.14, 1.22, and 1.23, and 1.07, 1.15, and 1.24, respectively. Analyzing fetal facial expression videos using AI may raise the possibility of being able to indirectly quantify brain activity. The ability to infer fetal brain activity via fetal facial expressions both qualitatively and quantitatively might be considered to have significant biological implications.

Keywords: 4D ultrasound; artificial intelligence; fetal brain function; fetal facial expression; free energy principle; chaotic dimension.

Introduction

Development of the fetal brain function is not yet fully understood. To understand brain function and the existence of consciousness, it is necessary to observe external output information from the brain[1]; therefore, at some point, the fetus may become conscious. Since there is no accurate method for observing electrical signals and metabolism in the brain from outside the body, muscle contractions caused by electrical signals from the brain may be considered representative of external output that can actually be observed. Facial expressions result from the integrated contraction of several groups of facial muscles; thus, it is considered reasonable to observe facial expressions to infer brain function. Advances in ultrasound technology have led to the widespread use of three-dimensional (3D) and four-dimensional (4D) ultrasound imaging to display fetal expressions in three-dimensions, making it possible to observe fetal expressions from outside the body. Various studies on fetal expressions using ultrasound have been conducted; however, in all cases, the ultrasound probe was placed on the mother's abdomen and fetal expressions were observed continuously, with the examiner recording only when changes in expression were noted [2,3,4,5]. Therefore, it was always difficult to recognize subtle facial expressions in a very short period and make diagnoses with minimal subjectivity. The potential development of a method to recognize facial expressions in a short period using a method with minimal subjective judgment may be useful for evaluating fetal brain function. In recent years, many artificial intelligence (AI) systems have been reported in the field of obstetrics and gynecology[6,7,8,9,10,11,12,13,14,15]. Using original AI that can classify fetal expressions by creating confidence scores for each of seven types of expressions on static images[16], we analyzed the collected fetal expressions and performed chaotic dimensional analysis, revealing that there are at least two different states of fetal expressions[17]. By interpreting this fluctuation using the free energy principle that is based on a variational Bayesian estimate to provide a comprehensive explanation of perception, action, emotion, sentiment, and decision-making[18,19,20,21,22], we quantitatively demonstrated the possibility of fetal brain activity.

We applied this AI to fetal videos and conducted expression analysis at 0.1-second intervals. This is a method of investigating fetal expressions qualitatively, quantitatively, and objectively. Previous studies demonstrated that chaotic dimensions could be calculated from fetal facial expressions and we, in this study, investigated the frequency, changes, and chaotic correlation dimensions during major expressions in fetal expression videos using AI, both qualitatively and quantitatively, and report our considerations as a hypothesis on the state of the fetal brain indicated by these expressions.

Materials and Methods

Acquisition of fetal facial expression data

The method for acquiring fetal facial expression data was detailed in our published paper (Y. Miyagi, 2022)[17]. Informed consent was obtained from all participants at Miyake Clinic between December 13, 2023, and February 21, 2024, with all data being anonymized. This retrospective, noninterventional study was performed in line with the principles of the Declaration of Helsinki and approved by the institutional review board of Miyake Clinic (Jan 22, 2024. No. mcg2024-1)[16,23].

Videos of fetal faces from consecutive normal singleton pregnancies at 28 to 37 weeks of gestation were recorded in MP4 format at 10.008 frames per second using 4D ultrasound with GE Voluson E10 BT20 (GE Healthcare, Zipf, Austria) and a curved array trans-abdominal transducer (GE eM6C G2, 2 - 7 MHz). These videos were transferred to an offline AI system[16] at Medical Data Labo, Japan. The development of the AI classifier with the original deep neural network architecture consisted of 13 layers; 2 convolution layers, 3 rectified linear unit layers, 2 pooling layers, 1 flatten layer, 3 linear layers, 1 batch normalization layer, and 1 softmax layer. The number of fetus/images were 237/1,457 and the number of test/validation/ training dataset for creating the AI with data augmentation, five-fold cross validation, L2-reglarization and early stopping was 251/1,536/11,248. The accuracy/sensitivity/specificity values were 0.996/0.964/1.000, 1.000/1.000/1.000, 0.996/1.000/0.994, 1.000/1.000/1.000, 1.000/1.000/1.000, 1.000/1.000/1.000, and 1.000/1.000/1.000 for eye blinking, mouthing, neutral, scowling, smiling, tongue expulsion, and yawning, respectively. Each video frame was converted into JPG-format images, cropped to 100 × 100 pixels, and divided by an AI classifier into seven confidence scores for each expression category such as eye blinking, neutral, mouthing, scowling, smiling, tongue expulsion, and yawning[16,17]. A seven-dimensional (7D) vector that consisted of confidence scores of the time-series per fetus was obtained:

$$\boldsymbol{x}_t = \{x_{t1},\ x_{t2}, \dots,\ x_{t7}\}^T$$

where $\boldsymbol{x}_t$: seven elements of fetal facial expressions at time t. The vector with the largest value is determined as the expression.

We applied the 7D vector to a practical algorithm to determine the character of strange attractors[24,25,26,27,28] to analyze multi-dimensional data. For a 7D vector of a time-series, we reconstructed the vector $\boldsymbol{x}_i$ by shifting time τ:

$$\boldsymbol{Y_j} = \{{}^{t}\boldsymbol{x}_{jk},\ {}^{t}\boldsymbol{x}_{j,k+\tau},\ \ldots,\ {}^{t}\boldsymbol{x}_{j,k+(\mathrm{m}-1)\,\tau}\},\ (k = 1, 2, \ldots, \mathrm{N}_i)$$

where τ is the time, j is the facial category number, m is the embedding dimension, and N is the number of video frames.

We then calculated the correlation dimension, D_2, as follows:

$$C_j(r) = \frac{1}{{N_j}^2}\sum_{f=1}^{N_i}\sum_{g=1}^{N_i} Q(r - |\boldsymbol{Y_{jf}} - \boldsymbol{Y_{jg}}|)$$

$$p_j = \frac{1}{N_{ij}}\sum_{g=1}^{N_j} Q(r - |\boldsymbol{Y_{jf}} - \boldsymbol{Y_{jg}}|)$$

$$D_{2j} = \lim_{r\to 0}\frac{\log C_j(r)}{\log\ r} = \lim_{r\to 0}\frac{\sum_{j=1}^{N_j} p_j^2}{\log r}$$

where r is $\{r \in \mathbb{R} \mid r > 0\}$ and Q is a Heaviside step function.

Changes in fetal facial expressions

We determined the spatial relationships of facial features from 922 images in seven categories used for AI created with 14,208 images from January 1, 2020, to September 30, 2020, (IRB No.: 2019-10)[23]. Then, quantitative transition patterns of facial expressions, focusing on the duration of each transition, were sought for a completely different video collected for this study between December 13, 2023, and February 21, 2024. Using our original AI for classifying fetal facial expressions reported in 2021[16], the last NetPort and softmax layer of AI were removed and facial features were extracted so that facial expressions could be placed in 2D and 3D space based on their relevance.

The number of expression data was unified to the minimum number of images obtained, then principal component analysis[29] was used for dimension reduction. The norm from the coordinate center was calculated, and a dimensional reduction method with no difference was selected to create 2D and 3D spaces, to which the video data obtained this time were applied.

Next, when a representative expression observed over a long period of time lasted for more than one second, the duration of that expression and expressions before and after it were analyzed. This is because the time required for the appearance of facial expressions from the recognition of facial expressions from others is a neuroscientific issue, and there does not appear to be an established time frame even in adults. Therefore, in this study, we assumed for convenience that stable fetal facial expressions last for more than one second. Types of facial expression transitions and time required for the transition were investigated.

Analysis of observed fetal facial expression time

We analyzed the representative expression and the total observation time by counting the frame length for each expression based on all confidence score information. For each image frame, the fetal expression that shows the maximum value among the 7 confidence scores was selected. When transition from facial expression A to facial expression B, the time of facial expression A is measured. For the facial expressions that were seen frequently, we examined the facial expressions before and after them.

Changes in correlation dimension of brain activity inferred from fetal facial expressions

We selected representative expressions that were observed for a relatively long time from all expressions, and when those expressions lasted for more than 25 seconds, corresponding to the minimum 250 time-series data-points required to calculate the correlation dimension in this study, we examined changes in the correlation dimension of the confidence score for the 25 seconds before and after the expression regardless of whether the confidence score of the original expression is included or not. We interpreted the results based on the free energy principle.

Free energy principle for fetal facial expressions

The free energy, F, in generating fetal expressions using the free energy principle is as follows:

$$F(\tilde{o},\mu) = D_{KL}[Q(\tilde{s},\tilde{u}|\mu)||P(\tilde{s},\tilde{u}|\tilde{o})] - \ln P(\tilde{o}|mdl)$$

$$\mu_t = arg\,min_{\mu} F(\{o_0,\dots,o_{t+1}\},\mu)$$

$$a_t = \arg min_{a} \sum_{\Omega} P(o_{t+1}|o_t,a)\,F(\{o_0,\dots,o_{t+1}\},\mu_t)$$

$$\boldsymbol{x} \in a$$

$$\tilde{o} = (o_1, o_2, \dots, o_t)$$

$$\tilde{s} = (s_1, s_2, \dots\ s_t)$$

$$\tilde{u} = (u_1, u_2, \dots\ u_t)$$

, where a is actions, D_{KL} is Kullback–Leibler divergence[30,31], *mdl* is a model, o_t is observations, P is generative density, Q is recognition density, s_t is hidden states, u is prediction of the result of causing an action[30,31], Ω is a set of observations, μ is sufficient statistics[30,31], and μ_t is perception.

When facial expressions of a fetus are focused,

$$\boldsymbol{x} \approx a$$

Statistical analysis

Wolfram Language and Mathematica 13.2 (Wolfram Research, Champaign, IL, United States) were used for all as well as statistical analyses, and we also used the Kruskal-Wallis test for multiple comparisons and the Mann–Whitney test for the two group comparisons. We set P <0.05 as significant.

Results

Acquisition of fetal facial expression data

There were 47 videos from the target patients, with an average age of 30.69 ± 5.27 years (mean ± standard deviation: SD), and the minimum and maximum ages were 22 and 39 years, respectively. Gestational age was 32.87 ± 3.05 weeks, with a minimum and maximum of 28 and 38 weeks, respectively. There were 24 primi- and 23 multiparous women, respectively, with 25 male and 22 female fetuses. The total observation time was 95.27 minutes, 57,208 frames. The recording times were 138.8 ± 56.8 (mean ± SD) seconds, 27.8, 210.2, 151.6 and 41.9 – 200.7 seconds for minimum, maximum, median and 5-95 %ile values, respectively.

Analysis of observed fetal facial expression time

The observed frames for eye blinking, neutral, mouthing, scowling, smiling, tongue expulsion, and yawning were 127, 2,706, 8,732, 1,052, 157, 43 and 618, 141, 7,380, 8,632, 563, 50, 28 and 1,023, 452, 13,908, 16,008, 570, 557, 237 and 2,374 in from 28 to 30+6, from 31 to 34+6 and from 35 to 37+6 weeks of gestations, respectively. The gestational age had no effect (P=0.8789). As shown in Figure 1 and Table 1, when all confidence scores were collected, neutral had the highest number of observations at 47 times, with 71.0 ± 52.3, 9.4 – 174.8 (mean ± SD, 5 – 95 %ile) seconds, followed by mouthing at 45 times with 53.3 ± 45.9, 0.7 – 127.3 seconds. There was a significant difference in variation between expressions (P = 3.47×10^{-16}). Among the expressions, neutral, mouthing, and both neutral and mouthing were significantly longer in duration (P <0.05).

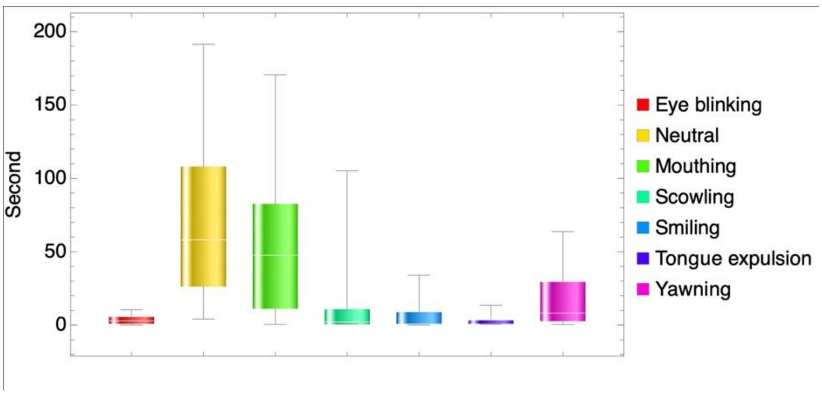

Figure 1: The total observation time for each facial expression. Neutral was the most common, with 71.0 ± 52.3, 9.4 – 174.8 (Mean ± Standard deviation, 5 – 95 %ile) seconds, followed by mouthing at 53.3 ± 45.9, 0.7 – 127.3 seconds. Significant differences were

observed in the variability between facial expressions (P =3.47×10^{-16}). Neutral, mouthing, and both neutral and mouthing were observed for significantly longer among facial expressions (P = 3.47×10^{-16}, P = 4.55×10^{-4}, P = 2.13×10^{-18}, respectively).

Table 1: Observation time (sec) of facial expressions detected from all videos by AI. There were significant differences among facial expression groups (P = 3.47×10^{-16}). Neutral, mouthing, and both neutral and mouthing were significantly longer than other facial expressions (P = 3.47×10^{-16}, P = 4.55×10^{-4}, P = 2.13×10^{-18}, respectively).

Facial expression	N	Mean	SD	Median	5 %ile	95 %ile
Eye blinking	20	3.60	2.95	2.85	0.1	8.1
Neutral	47	71.00	52.29	58.1	9.4	174.8
Mouthing	45	53.32	45.96	47.6	0.7	127.3
Scowling	15	14.57	28.19	2.10	0.7	105.2
Smiling	13	5.88	9.67	0.80	0.1	34.0
Tongue expulsion	11	2.80	4.02	0.70	0.6	13.5
Yawning	24	16.73	18.35	8.2	0.7	50.1

SD: Standard deviation.

Changes in fetal facial expressions

Table 2: Types of facial expression transitions and time required for the transition. The total observation time was 95.3 min. There were 36 different transitional patterns. The observed facial expressions lasted for 158.79 ± 434.56, 0.7 – 1469.7 (Mean ± SD, 5 – 95 %ile) seconds, with the longest and shortest being 2235.5 and 0.7 seconds, respectively. Throughout the entire period, there were an average of 9.97 changes in facial expressions. There was a significant difference between the observed times in 36 groups (P=2.64×10^{-16}).

From	To	n	Time (sec)	%Time	Mean	SD	Median	5%ile	95%ile
Eye blinking	Mouthing	14	31.1	0.54	2.22	2.07	1.40	0.10	6.90
Eye blinking	Neutral	13	18.4	0.32	1.42	0.89	1.10	0.60	3.50
Eye blinking	Scowling	2	6	0.10	3.00	3.25	3.00	0.70	5.30
Eye blinking	Smiling	2	5.2	0.09	2.60	1.70	2.60	1.40	3.80
Eye blinking	Tongue expulsion	1	0.7	0.01	0.70	NA	0.70	0.70	0.70
Eye blinking	Yawning	6	9.9	0.17	1.65	1.03	1.70	0.40	2.90
Mouthing	Eye blinking	14	192.6	3.37	13.76	22.15	4.30	0.60	77.70
Mouthing	Neutral	44	1469.7	25.71	33.40	35.81	14.70	0.70	122.40
Mouthing	Scowling	11	127.8	2.24	11.62	11.98	10.40	0.10	32.20
Mouthing	Smiling	8	94.6	1.65	11.83	14.08	4.25	0.30	34.80
Mouthing	Tongue expulsion	7	27.7	0.48	3.96	5.06	1.50	0.70	14.80
Mouthing	Yawning	19	379.4	6.64	19.97	25.60	9.30	0.10	86.70
Neutral	Eye blinking	13	70.8	1.24	5.45	7.36	3.80	0.60	28.30
Neutral	Mouthing	44	2237.5	39.14	50.85	40.28	42.70	4.60	115.80
Neutral	Scowling	10	69.7	1.22	6.97	13.85	1.40	0.10	44.30
Neutral	Smiling	7	48.3	0.84	6.90	7.82	3.60	0.70	22.30
Neutral	Tongue expulsion	4	17.1	0.30	4.28	4.65	2.60	0.80	11.10

Neutral	Yawning	22	194.3	3.40	8.83	7.75	7.60	0.70	24.60
Scowling	Eye blinking	2	11.1	0.19	5.55	6.72	5.55	0.80	10.30
Scowling	Mouthing	14	119.4	2.09	8.53	13.81	1.55	0.70	41.20
Scowling	Neutral	4	67	1.17	16.75	28.78	2.85	1.40	59.90
Scowling	Tongue expulsion	2	4.8	0.08	2.40	2.40	2.40	0.70	4.10
Scowling	Yawning	3	15.8	0.28	5.27	4.09	6.50	0.70	8.60
Smiling	Eye blinking	3	4.6	0.08	1.53	0.67	1.70	0.80	2.10
Smiling	Mouthing	10	39.9	0.70	3.99	6.36	1.40	0.10	21.20
Smiling	Neutral	5	17.8	0.31	3.56	4.15	1.40	0.60	10.40
Smiling	Yawning	4	9.9	0.17	2.48	3.17	1.05	0.60	7.20
Tongue expulsion	Mouthing	6	11.1	0.19	1.85	1.80	1.05	0.70	5.30
Tongue expulsion	Neutral	7	16.1	0.28	2.30	2.74	1.30	0.60	8.20
Tongue expulsion	Scowling	2	2.9	0.05	1.45	1.06	1.45	0.70	2.20
Tongue expulsion	Yawning	1	0.7	0.01	0.70	NA	0.70	0.70	0.70
Yawning	Eye blinking	5	11.6	0.20	2.32	3.44	1.30	0.10	8.40
Yawning	Mouthing	20	243.2	4.25	12.16	13.58	7.70	0.50	38.70
Yawning	Neutral	21	107.8	1.89	5.13	6.01	3.90	0.60	20.90
Yawning	Scowling	4	12	0.21	3.00	3.73	1.40	0.70	8.50
Yawning	Smiling	5	19.8	0.35	3.96	2.65	4.00	1.40	7.50

SD: Standard deviation

As shown in Table 2, there were 36 patterns of transition between facial expressions. The duration of the observed facial expressions was 158.79 ± 434.56, 0.7 – 1469.7 (Mean ± SD, 5 – 95 %ile) seconds, with the longest being 2,237.5 seconds and shortest at 0.7 seconds. Facial expression transitions occurred on 9.97 ± 10.26, 1 – 44 times. There were significant differences in the observed time intervals among the 36 groups ($P = 2.64 \times 10^{-16}$). There were 42 possible variations in facial expressions, and the following six were not noted: tongue expulsion to smiling, tongue expulsion to eye blinking, scowling to smiling, smiling to tongue expulsion, smiling to scowling, and yawning to tongue expulsion. The total of duration of facial expression before the expression change accounted for 2,237.5 seconds (39.14%) of the total observation time of 5,716.3 seconds, with the longest transition being from neutral to mouthing, followed by mouthing to neutral at 1,469.7 seconds (25.71%). Combined, these two transitions accounted for 64.85% of the total observation time.

Placing facial expressions in space using dimensional reduction methods, we chose principal component analysis, which did not significantly differ in the norm of each facial expression. The norm values (mean ± SD, 5 – 95 %ile) for eye blinking, neutral, mouthing, scowling, smiling, tongue expulsion, and yawning were 10.92 ± 2.17, 9.06 – 14.63, 11.42 ± 1.74, 8.96 – 13.53, 12.04 ± 1.43, 10.55 – 14.27, 9.87 ± 1.59, 7.82 – 12.37, 13.05 ± 1.11, 11.95 – 14.63, 10.90 ± 1.94, 7.82 – 12.75, and 12.17 ± 1.47, 10.45 – 13.85 in 2D space, and 60.52 ± 5.08, 53.12 – 66.63, 64.49 ± 9.39, 56.13 – 83.03, 64.54 ± 9.08, 56.25 – 80.24, 60.14 ± 9.17, 52.07 – 75.90, 61.81 ± 6.80, 52.07 – 70.00, 62.15 ± 8.19, 55.64 – 75.90, and 71.76 ± 7.72, 66.06 – 83.03 in 3D space, respectively. Although there was no significant difference, the norm value of yawning was the largest in both 2D and 3D spaces. Figure 2 shows how the coordinates for each expression were established in 2D and 3D spaces. The coordinates of neutral and mouthing expressions were close to each other in both spaces, with frequent transitions, especially from neutral to mouthing and vice versa.

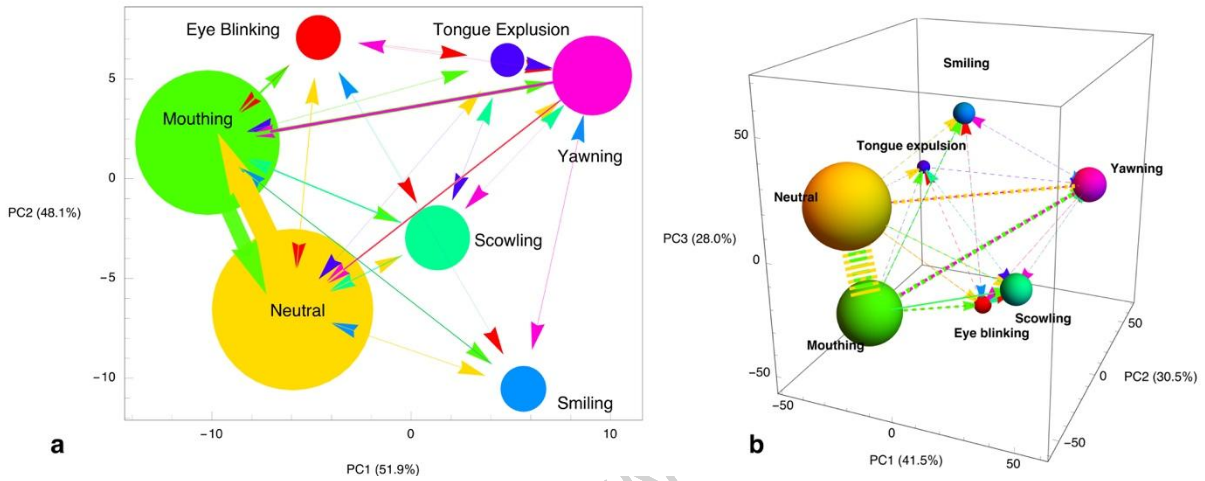

Figure 2: Facial expressions placed in 2D and 3D spaces and facial expression transitions. This figure shows how the coordinates for each expression were established in 2D (a) and 3D (b) spaces, using data from 922 labeled images of fetal expressions.

We applied a principal component analysis method that showed no significant differences in norm from the coordinate center. Video data from this study were then integrated into these spatial models for improved visualization of facial expression relationships. The size of each circle or sphere indicates the observation duration, while arrow colors correspond to the initial expression, with arrow diameters reflecting the transition frequency. Neutral and mouthing expressions are close to each other, with frequent transitions, especially from neutral to mouthing and vice versa.

As shown in Table 3, neutral expressions lasting for more than one second showed significant differences in duration before and after neutral (P =0.00004 and P =0.00002, respectively). Mouthing before neutral lasted for 16.40 ±16.49, 0.4 – 54.5 (Mean ± SD, 5 – 95 %ile) seconds (87.3% of total duration before neutral) and mouthing after neutral lasted for 13.49 ± 18.56, 0.1 – 64.9 seconds (90.5% of total duration after neutral). Neutral was reached 5.72 ± 6.45, 1.26 – 16.39 seconds after some expressions, and the next expression transitioned after 3.47

± 4.95, 0.64 – 13.49 seconds (P =0.296). For mouthing lasting for one second or longer, there were significant differences in duration before and after (P =9.57×10^{-6}). Neutral before mouthing transitioned to mouthing after 11.43 ± 11.49, 0.7 – 35 seconds, and neutral after mouthing averaged 17.63 ± 16.09, 0.8 – 53.3 seconds. Mouthing occurred 5.04 ± 3.88, 1.2 – 11.42 seconds after the preceding expression, and the next expression transitioned after 5.33 ± 6.33, 1.23 – 17.62 seconds (N.S.).

Table 3: Before-and-after expressions and their duration when neutral (upper panel) and mouthing (lower panel) facial expressions lasted for more than 1 second. There was a significant difference between the time required for before and after neutral (P =0.00004 and P =0.00002, respectively). Mouthing, which preceded neutral, took an average of 16.40 ±16.49, 0.4 – 54.5 (Mean ± SD, 5 – 95 %ile) seconds (87.3% of the total duration of facial expressions before neutral) before transitioning to neutral. The average post-neutral mouthing time was 13.49 ± 18.56, 0.1 – 64.9 seconds (90.5% of the total duration of facial expressions after neutral). On average, after 5.72 ± 6.45, 1.26 – 16.39 seconds of some facial expressions, it became neutral, and the next expression after neutral transitioned in 3.47 ± 4.95, 0.64 – 13.49 seconds (P = 0.296).

There was a significant difference between the time required for facial expressions before and after mouthing that lasted for more than 1 second (P = 9.57×10^{-6}). The average neutral time after mouthing was 17.63 ± 16.09, 0.8 – 53.3 seconds. The leading expression became mouthing after 5.04 ± 3.88, 1.20 – 11.42 seconds, and the next expression after mouthing transitioned after 5.33 ± 6.33, 1.23 – 17.62 seconds (N.S.).

	Eye blinking	Mouthing	Scowling	Smiling	Tongue expulsion	Yawning
Before-neutral						
Mean	1.34	16.40	11.13	2.13	1.26	2.08
SD	0.82	16.50	15.60	1.90	1.12	1.47
Median	1.40	11.20	3.20	1.60	0.70	1.40
5 %ile	0.4	0.4	1.1	0.6	0.3	0.1
95 %ile	2.8	54.5	29.1	4.7	3.1	4.7
After-neutral Mean	0.69	13.49	2.20	1.90	0.64	1.91
SD	0.93	18.56	2.70	2.43	0.87	1.36
Median	0.40	6.40	0.90	1.25	0.40	2.00
5 %ile	0	0.1	0.4	0	0	0
95 %ile	2.7	64.9	5.3	5.1	2.1	4.3

	Eye blinking	Neutral	Scowling	Smiling	Tongue expulsion	Yawning
Before-mouthing Mean	2.73	11.43	5.79	1.97	1.20	7.17
SD	1.52	11.49	8.57	2.17	0.63	6.34
Median	2.30	8.00	0.80	0.70	1.0	6.30
5 %ile	0.7	0.7	0.1	0.3	0.7	0.8
95 %ile	5.0	35.0	24.9	5.8	2.2	21.3
After-mouthing Mean	2.37	17.63	1.23	2.81	1.32	6.61
SD	1.81	16.10	1.85	4.69	1.47	9.10
Median	3.00	15.30	0.70	0.70	0.70	2.10

5 %ile	0.0	0.8	0.0	0.0	0.0	0.0
95 %ile	4.7	53.3	5.7	12.7	3.8	31.4

SD: Standard deviation.

Changes in correlation dimension of brain activity inferred from fetal facial expressions

Table 4: Correlation dimension before neutral, during neutral, after neutral, and before mouthing, during mouthing, and after mouthing facial expressions. There was no significant difference among the periods for each expression. There was also no significant difference between neutral and mouthing.

	Before Neutral	During Neutral	After Neutral	Before Mouthing	During Mouthing	After Mouthing
Mean	1.15	1.21	1.20	1.06	1.16	1.16
SD	0.16	0.25	0.30	0.28	0.22	0.26
Median	1.14	1.22	1.23	1.08	1.15	1.24
5%ile	0.99	0.67	0.83	0.70	0.83	0.75
95%ile	1.42	1.52	1.52	1.38	1.52	1.41

SD: Standard Deviation.

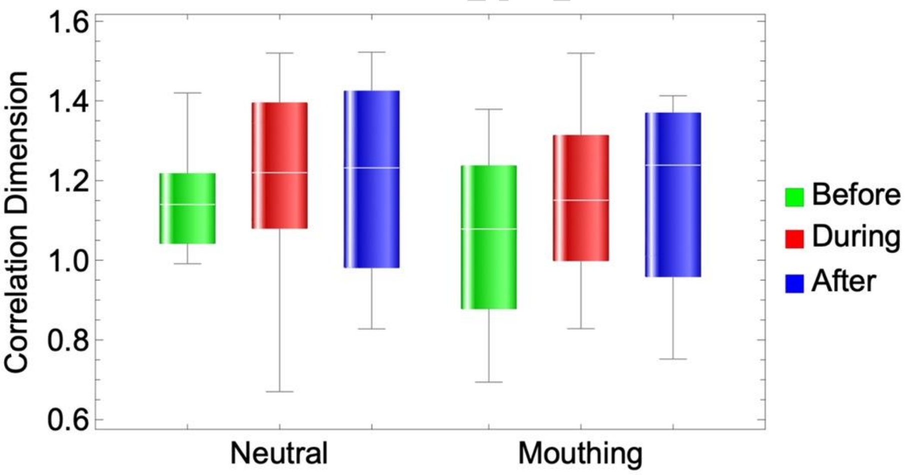

Figure 3: Correlation dimension of confidence score for each 25 seconds before and after neutral and mouthing sustained for more than 25 seconds. There was no intergroup variation in either neutral or mouthing. Focusing on the median value of the correlation dimension, there was no significant difference; but in neutral, pre-neutral (1.14; median value) was less than during neutral (1.22); post-neutral (1.23) was almost the same as during neutral; in mouthing, pre-mouthing (1.07) was less than during mouthing (1.15); post-mouthing (1.24) increased from mouthing; and the values post-mouthing, during-neutral, and post-neutral tended to be almost the same.

As shown in Table 4 and Figure 3, the correlation dimension of the confidence score for each 25 seconds before and after neutral and mouthing sustained for more than 25 seconds was calculated. The correlation dimensions for before, during, and after neutral were 1.15 ± 0.16, 0.99 – 1.42, 1.21 ± 0.25, 0.67 – 1.52, and 1.20 ± 0.30, 0.82 -1.52 (mean ± SD, 5 – 95 %ile), respectively. The correlation dimensions for before, during, and after mouthing were 1.06 ± 0.28, 0.69 – 1.37, 1.16 ± 0.22, 0.82 – 1.51, and 1.16 ± 0.25, 0.75 – 1.41, respectively. There were no significant differences between groups for either neutral or mouthing. There was no significant difference in the correlation dimensions between neutral and mouthing.

Discussion

Using AI for fetal facial expression recognition, it has now become possible to analyze fetal facial expressions in detail on a frame-by-frame basis from ultrasound videos. In previous studies, much time and effort were required, and the determination of events was based on subjective judgment. Here, by quantitatively analyzing a total of 95.27 minutes, 57,208 frames, of video on a frame-by-frame basis and in single-frame increments, we were able to clarify the characteristics of fetal facial expressions. Based on frequency, neutral and mouthing were significantly more common, so both were considered important information. Mouthing is frequently observed in studies and considered an important expression[2]. It was significantly more frequent than other facial expressions early in the third trimester[4], being consistent with our results. We consider that this study provides important new insights by quantitatively demonstrating for the first time the possibility that neutral expressions have meaning, which has not been emphasized to date.

The gestational age had no effect for the observed time for facial expression profiles. Figure 1 and Table 1 show significant differences between facial expressions, suggesting that fetal facial expressions have meaning beyond reflexes. Table 2 also presents significant differences, with six of the 42 facial expression transitions not observed, indicating that facial expression manifestation is not random but suggests some brain function. Furthermore, there were mutual changes in expressions between neutral and mouthing, with transitions between the two accounting for 64.85%.

As the method of dimensional reduction based on our AI, we selected principal component analysis; a statistical method used to reduce the dimensionality of data by transforming it into a new set of orthogonal axes that capture the most variance[29]. The flow map showed that neutral and mouthing were close to each other in both 2D and 3D spaces (Figure 2). Furthermore, there was a significant difference in the duration of expressions before and after neutral and mouthing that lasted for more than one second (Table 3). Therefore, it was considered that neutral and mouthing were the basic states of fetal facial expressions.

The correlation dimension of confidence scores for each 25 seconds before and after neutral and mouthing sustained for more than 25 seconds suggested that there would be some brain activities when facial expression changes. We proposed the hypothesis that when the chaotic dimension of the 7D time series vector created by AI from fetal facial expressions is large, brain activity and free energy are both high, and when the dimension is low, that are consequently both low[30,31]. The free energy principle with active inference is a theory explaining cognition and brain behavior. It uses variational Bayesian estimates to describe perception, action, emotion, sentiment, and decision-making. To maintain equilibrium, an agent minimizes informational free energy and prediction error. This involves adjusting internal states and environmental sampling to reduce free energy[18,19,20,21,22]. Biologically, the relationship between chaotic dimensions such as correlation dimensions derived from fetal expressions and brain activity has not been proven. Although no significant differences were noted in this study, Figure 3 might suggest the existence of fluctuations in correlation dimensions. In living organisms, brain activity optimizes and minimizes energy consumption for survival, and only very slight dimensional changes might be able to be observed, and there might be no difference significant enough to reach the typical biological significance level of α error = 0.05. There are no reports in fetuses regarding of this hypothesis. However, we speculated that the fluctuations in the correlation dimension obtained from fetal facial expressions might be a clue to access the fetal brain activity. Neutral may have high free energy, similar to after mouthing. Incidentally, while we attempted to infer brain function using ultrasound, there are reports of detecting brain damage in rabbit fetuses via MRI as a model for cerebral palsy[32].

This pilot study demonstrated that AI could enable quantitative observation of fetal facial expressions using non-invasive ultrasound. Regarding the application of AI using dynamic and static ultrasound, there are studies comparing AI with diagnoses made by ultrasonographers and cardiologists in heart failure patients[33]. However, our current study did not compare human versus AI performance; instead, we observed outputs from AI that humans cannot produce. As seen in rats with severe uncontrolled maternal hyperglycemia leading to neurodevelopmental delay, abnormal conditions such as maternal hyperglycemia in humans may also affect the fetal brain, causing the fetus to develop pathological conditions and exhibit characteristic changes in facial expressions[34]. Since the data did not compare normal fetuses with pathological fetuses in this pilot study, no normal values were established. However, accumulating data in the future may enable determination of what constitutes normal. Observing fetal facial expressions during ultrasound examinations is easily feasible in routine clinical practice. Recording a few minutes of video allows obtaining the frequency and duration of each expression, as well as chaotic dimensions, for individual fetuses. Comparing these to future normal values could enable clinical classification based on facial expressions. Since facial expressions are thought to correlate with brain function, this research method may indirectly assess fetal brain development.

Further studies with longitudinal design of increased sample size would be needed. Furthermore, if methods for evaluating the neurophysiological functions of the fetal brain are established in the future and their relationship with fetal facial expressions is clarified, the meaning behind changes in fetal facial expressions may become clearer. However, evaluating brain function requires caution due to ethical and clinical implications. Diagnosis and interpretation must consider various confounding variables, potentially necessitating consensus involving not only physicians but also ethicists or the public.

Limitations

As for limitations, firstly, the details of fetal facial muscle contraction are unclear. Current diagnostic devices cannot resolve muscle contractions shorter than 0.1 seconds. The refractory period of human skeletal muscle is approximately 1–5 msec[35], and imaging at shorter intervals would enable observation of detailed muscle contraction fluctuations. In the future, if the frame rate of 4D ultrasound diagnostic devices increases to 1,000 frames per second (100 times the current rate), it would then be possible to perform precise physiological analysis. Secondly, approximately 250 images seemed to require calculating the chaotic dimension. Therefore, the frame rate is important, and currently it takes about 25 seconds (250 images). If there is a device with a high frame rate, this time can be shortened. Using a device with a frame rate 100 times higher than the current one, it will be possible to calculate the dimension of facial expressions in 0.25 seconds. However, regarding the statistical analysis of chaotic dimensional fluctuations, if the number of cases increases, it is expected that a significant difference will be observed even with the current device. Third, since this analysis combines all cases between 28 and 38 weeks, changes in development based on week classification are unclear. However, if we analyze the data by week classification, we may be able to see changes in development. Fourth, we used principal component analysis for the spatial flow map of facial expressions, but depending on the dimension reduction method used, neutral and mouthing may not be located close to each other, and the flow map may vary depending on the AI. Fifth, the relationship between REM and non-REM states of the brain and fetal facial expressions remains unclear. Although REM has been reported in fetuses[36], until a method capable of simultaneously collecting facial expression and REM information is developed, this relationship will remain unclear. Sixth, sample size might not be large enough, resulting in possible limitation statistical reliability.

Conclusions

Development of the fetal brain remains largely unknown. However, analyzing fetal facial expression videos using AI suggests the possibility of being able to indirectly quantify brain activity. Even indirectly, inferring fetal brain activity qualitatively and quantitatively would be considered to have significant biological implications. In the future, we hope to analyze information on brain activity in various environments to enable intrauterine diagnosis of fetal stress. The methodology demonstrated in this study might provide clues for delivering appropriate care to improve the fetal environment.

Data availability statement

The datasets generated during this study are available from the corresponding author upon reasonable request.

References

1. Lagercrantz, H. The birth of consciousness. *Early Hum. Dev.* **85** (10 Suppl), S57-58 (2009).
2. Kanenishi, K., Hanaoka, U., Noguchi, J., Marumo, G. & Hata, T. 4D ultrasound evaluation of fetal facial expressions during the latter stages of the second trimester. *Int. J. Gynaeco.l Obstet.* **121**, 257-260 (2013).
3. Sato, M. et al. 4D ultrasound study of fetal facial expressions at 20-24 weeks of gestation. *Int. J. Gynaecol. Obstet.* **126**, 275-279 (2014).
4. Yan, F. et al. Four-dimensional sonographic assessment of fetal facial expression early in the third trimester. *Int. J. Gynaecol. Obstet.* **94**, 108-113 (2006).
5. Reissland, N., Francis, B., Mason, J. & Lincoln, K. Do facial expressions develop before birth?. *PLoS One* **6**, e24081 (2011).
6. Miyagi, Y., Habara, T., Hirata, R. & Hayashi, N. Predicting implantation by using dual AI system incorporating three-dimensional blastocyst image and conventional embryo evaluation parameters—A pilot study. *Reprod. Med. Biol.* **23**, e12612 (2024).
7. Miyagi, Y., Habara, T., Hirata, R. & Hayashi, N. Deep Learning to predicting live births and aneuploid miscarriages from images of blastocysts combined with maternal age. *Int. J. Bioinfor. Intell. Comput*. **1**, 10-21 (2022).
8. Miyagi, Y., Habara, T., Hirata, R. & Hayashi, N. Predicting a live birth by artificial intelligence incorporating both the blastocyst image and conventional embryo evaluation parameters. *Artif. Intell. Med. Imaging.* **1**, 94-107 (2020).
9. Miyagi, Y., Habara, T., Hirata, R. & Hayashi, N. Feasibility of artificial intelligence for predicting live birth without aneuploidy from a blastocyst image. *Reprod. Med. Biol.* **18**, 204-211 (2019)
10. Miyagi, Y., Habara, T., Hirata, R. & Hayashi, N. Feasibility of predicting live birth by combining conventional embryo evaluation with artificial intelligence applied to a blastocyst image in patients classified by age. *Reprod. Med. Biol.* **18**,344–356 (2019).
11. Miyagi, Y. & Miyake, T. Potential of artificial intelligence for estimating Japanese fetal weights. *Acta Medica Okayama* **74**, 483-93 (2020).
12. Miyagi, Y. et al. A novel method for determining fibrin/fibrinogen degradation products and fibrinogen threshold criteria via artificial intelligence in massive hemorrhage during delivery with hematuria. *J. Clin. Med.* **13**, 1826 (2024).
13. Miyagi, Y. et al. New method for determining fibrinogen and FDP threshold criteria by artificial intelligence in cases of massive hemorrhage during delivery. *J. Obstet. Gynaecol. Res.* **46**, 256–265 (2020).

14. Miyagi, Y., Fujiwara, K., Nomura, H., Yamamoto, K. & Coleman. R. L. Feasibility of new method for the prediction of clinical trial results using compressive sensing of artificial intelligence. *Br. J. Health. Med. Res*. **10**, 237-267 (2023).
15. Miyagi, Y., Takehara, K., Nagayasu, Y. & Miyake, T. Application of deep learning to the classification of uterine cervical squamous epithelial lesion from colposcopy images combined with HPV types. *Oncol. Lett.* **19**, 1602-1610 (2020).
16. Miyagi, Y., Hata, T., Bouno, S., Koyanagi, A. & Miyake, T. Recognition of fetal facial expressions using artificial intelligence Deep Learning. *Donald School J. Ultrasound Obstet. Gynecol.* **15**, 223-228 (2021).
17. Miyagi, Y., Hata, T., Bouno, S., Koyanagi, A. & Miyake, T. Artificial intelligence to understand fluctuation of fetal brain activity by recognizing facial expressions. *Int. J. Gynecol. Obstet.* **161**, 877-885 (2023).
18. Friston, K., Kilner, J. & Harrison, L. A free energy principle for the brain. *J. Physiol. Paris.* **100**, 70–87 (2006).
19. Friston, K. The free-energy principle: a rough guide to the brain? Trends *Cognit. Sci.* **13**, 293–301 (2009).
20. Friston, K., Daunizeau, J., Kilner, J. & Kiebel, S. J. Action and behavior: a free-energy formulation. *Biol. Cybern.* **102**, 227–260 (2010).
21. Friston, K., Daunizeau, J. & Kiebel, S. J. Reinforcement learning or active inference? *PLoS One* **4**, e6421 (2009).
22. Miyagi, Y. et al. Kinetic energy and the free energy principle in the birth of human life. *Reprod. Med.* **5**, 65-80 (2024).
23. Miyagi, Y., Hata, T., Bouno, S., Koyanagi, A. & Miyake, T. Recognition of facial expression of fetuses by artificial intelligence (AI). *J. Perinat. Med.* **49**, 596-603 (2021).
24. Miyagi, Y., Miyagi, Y., Terada, S. & Kudo, T. Variations of multifractal structure in the fetal heartbeats. *Acta Med Okayama* **57**, 49–52 (2003).
25. Grassberger, P. & Procaccia, I. Characterization of strange attractors. *Phys. Rev. Lett.* **50**, 346–349 (1983).
26. Grassberger, P. Dimensions and entropies of strange attractors from a fluctuating dynamics approach. *Physica D* **13**, 34–54 (1984).
27. Farmer, J. D., Ott, E. & York, J. A. The dimension of chaotic attractors. *Physica D* **7**, 153–80 (1983).
28. Halsey, T. C., Jensen, M. H, Kadanoff, L.P., Procaccia, I. & Shraiman. B. I. Fractal measures and their singularities: The characterization of strangesets. *Phys. Rev. A* **33**, 1141–1151

(1986).

29. Pearson, K. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine* **2**, 559-572 (1901).
30. Miyagi, Y., Hata, T. & Miyake, T. Fetal brain activity and the free energy principle. *J. Perinat. Med*. **51**, 925-931 (2023).
31. Miyagi, Y., Hata, T. & Miyake, T. Fetal brain function and artificial intelligence in *Donald school textbook of ultrasound in obstetrics and gynecology, 5th ed*. (eds. Kurjak, A. & Chervenak, F. A.) 722-740 (Jaypee Brothers Medical Publishers Ltd, 2025).
32. Ambwani, G., Shi, Z., Luo, K., Jeong, J. W. & Tan, S. Distinguishing Laterality in Brain Injury in Rabbit Fetal Magnetic Resonance Imaging Using Novel Volume Rendering Techniques. *Dev. Neurosci*. **47**, 55-67 (2025). Erratum in: *Dev. Neurosci*. **47**, 420 (2025).
33. Liu, Z. et al. A generalized deep learning model for heart failure diagnosis using dynamic and static ultrasound. *J. Transl. In.t Med*. **11**, 138-144 (2023).
34. Piazza, F. V., et al. Severe Uncontrolled Maternal Hyperglycemia Induces Microsomia and Neurodevelopment Delay Accompanied by Apoptosis, Cellular Survival, and Neuroinflammatory Deregulation in Rat Offspring Hippocampus. *Cell Mol Neurobiol.* **39**, 401–414 (2019).
35. Dulhunty, A. F. A refractory period after brief activation of mammalian skeletal muscle fibers. *Neuroscience Lett.* **14**, 223–228 (1979).
36. Horimoto, N., Koyanagi, T., Nagata, S., Nakahara, H. & Nakano, H. Concurrence of mouthing movement and rapid eye movement/non-rapid eye movement phases with advance in gestation of the human fetus. *Am. J. Obstet. Gynecol.* **161**, 344-351 (1989).

Funding
None

Author contributions:
All authors have accepted responsibility for the entire content of this manuscript and approved its submission. The roles of the authors were as follows. Yasunari Miyagi: Conceptualization, methodology, software, formal analysis, investigation, writing original draft, reviewing and editing, visualization, supervision. Toshiyuki Hata: Validation, resources, data curation, review and editing, supervision, project administration. Takahito Miyake: Validation, resources, data curation, review and editing, supervision, project administration.

Ethics declarations

Competing interests statement

The authors declare no competing interests.

Generative AI and AI-assisted technologies in the writing process

No generative AI or AI-assisted technologies were used. The authors thoroughly reviewed and made additional corrections to ensure accuracy and appropriateness. The authors take full responsibility for the final content and conclusions presented in this publication.

Consent to participate

Informed consent was obtained from all individual participants included in the study.