European Heart Journal - Digital Health (2025) **6**, 683–687
https://doi.org/10.1093/ehjdh/ztaf039

ORIGINAL ARTICLE

Synthetic electrocardiograms for Brugada syndrome: from data generation to expert cardiologists evaluation

Beatrice Zanchi [1,2], Giuliana Monachino [1,3], Francesca Dalia Faraci [1], Matteo Metaldi [1], Pedro Brugada [4], Georgia Sarquella-Brugada [5,6,7], Elijah R. Behr [8], Josep Brugada [9], Lia Crotti [10,11], Bernard Belhassen [12,13], and Giulio Conte [14,15,]*

[1]Department of Innovative Technologies, Institute of Digital Technologies for Personalized Healthcare of SUPSI, Lugano, Switzerland; [2]Department of Quantitative Biomedicine, University of Zurich, Zurich, Switzerland; [3]Institute of Informatics, University of Bern, Bern, Switzerland; [4]Heart Rhythm Management Centre, Postgraduate Program in Cardiac Electrophysiology and Pacing, Universitair Ziekenhuis Brussel—Vrije Universiteit Brussel, European Reference Networks Guard-Heart, Brussels, Belgium; [5]Departament de Ciencies Mediques, Facultat de Medicina, Universitat de Girona, Girona, Spain; [6]Arrhythmia, Inherited Cardiac Diseases and Sudden Death Unit, Hospital Sant Joan de Déu, Esplugues, Barcelona, Spain; [7]Arrítmies Pediatriques, Cardiologia Genetica I Mort Sobtada, Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain; [8]Cardiovascular and Genomics Research Institute, City St George's, University of London and Cardiovascular Clinical Academic Group, St George's University Hospitals NHS Foundation Trust, London, United Kingdom; [9]Cardiology Department, Arrhythmia Section, Hospital Clinic, Universitat de Barcelona, Barcelona, Spain; [10]Center for Cardiac Arrhythmias of Genetic Origin and Laboratory of Cardiovascular Genetics, Istituto Auxologico Italiano, IRCCS, Milan, Italy; [11]Departement of Medicine and Surgery, University Milano Bicocca, Milan, Italy; [12]Hadassah Medical Center, Heart Institute, Kyriat Hadassah, Jerusalem, Israel; [13]Sackler School of Medicine, Tel-Aviv University, Tel-Aviv, Israel; [14]Division of Cardiology, Cardiocentro Ticino Institute, Ente Ospedaliero Cantonale, Via Tesserete 48, Lugano CH-6900, Switzerland; and [15]Faculty of Biomedical Sciences, Università della Svizzera Italiana USI, Lugano, Switzerland

Received 10 January 2025; revised 21 February 2025; accepted 13 March 2025; online publish-ahead-of-print 24 April 2025

Aims

Synthetic electrocardiograms (ECGs) for inherited cardiac diseases may overcome the issue related to data scarcity for artificial intelligence (AI)-based algorithms. This study aimed to evaluate experienced cardiologists' ability to differentiate synthetic and real Brugada ECGs.

Methods and results

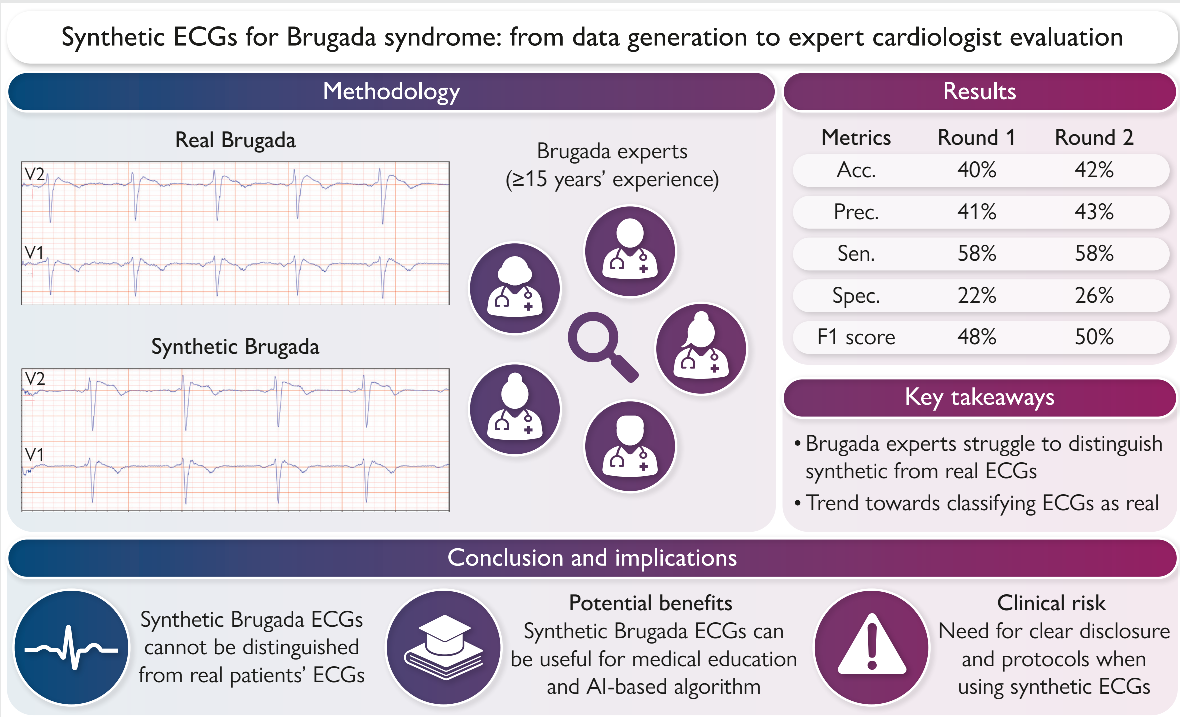
A total of 2244 ECG instances (50% synthetic generated by a generative adversarial network, 50% real Brugada patients' ECGs) were evaluated by 7 cardiologists, each with >15 years of experience. All ECGs were standard 12-lead recordings acquired with identical settings (paper speed 25 mm/s, amplitude 10 mm/mV) and randomly assigned without identifying markers. The examination was blinded and conducted in 2 rounds with at least 2 h gap between rounds to assess potential learning effects and intra-rater reliability. Each physician classified the recordings as 'real' or 'synthetic' without having any additional information. Performance metrics, including accuracy, sensitivity, specificity, and intra-rater reliability (Cohen's Kappa), were analyzed. Brugada syndrome (BrS) specialists' repeated evaluations were characterized by low accuracy (first round 40%, second round 42%), specificity (first round 22%, second round 26%) and sensitivity (first round 58%, second round 58%). Intra-rater reliability varied widely (Cohen's Kappa: −0.12 to 0.80).

Conclusion

Synthetic Brugada ECGs cannot be adequately distinguished from real patients' ECGs by BrS specialists.

* Corresponding author. Tel: +41 (0)91 811 53 63, Email: giulio.conte@eoc.ch

Graphical Abstract



Keywords Brugada Syndrome • Synthetic ECG • Artificial Intelligence • Machine learning • AI-enabled ECG • Cardiogenetics

Introduction

Brugada syndrome (BrS) is an inherited arrhythmia syndrome characterized by distinctive electrocardiographic (ECG) patterns and an increased risk of sudden cardiac death.[1,2] Despite established diagnostic criteria, the identification of BrS remains challenging due to the dynamic nature of ECG manifestations, phenotypic variability, and the need for an expert interpretation.

The emergence of artificial intelligence (AI) and machine learning (ML) approaches in cardiology has shown promise in improving diagnostic accuracy and efficiency.[3–10] However, the development and validation of AI-enabled ECG diagnostic tools for rare cardiac diseases like BrS faces a significant obstacle: the scarcity of available training data to build comprehensive datasets for AI-model development.

Synthetic ECG generation, powered by advanced ML techniques, has emerged as a potential solution to overcome data scarcity.[11,12] Synthetic ECGs could potentially augment existing datasets for AI training, enhance medical education, and help the development of more robust diagnostic algorithms.[13,14] However, to be valuable for clinical applications and research, synthetic ECGs must faithfully reproduce the subtle characteristics and variations seen in real patient recordings. This requirement raises a critical question: can synthetic ECGs be distinguished from genuine clinical data, even by expert clinicians? This study addresses this fundamental question by investigating experienced cardiologists' ability to differentiate synthetic and real Brugada ECGs. Signal's visual inspection represents the most widely adopted approach to evaluate the quality of synthetic datasets, with other metrics such as statistical distribution matching and time-domain analysis being less commonly employed.[13,14] Our aims were to: (i) assess the distinguishability of synthetic BrS ECGs from real patient recordings through expert evaluation, and (ii) analyse the consistency of expert judgements across multiple assessments. Expected clinical implications of this study rely on the potential of validated synthetic ECGs to serve as an educational resource for healthcare professionals at all levels, enabling comprehensive training in BrS pattern recognition despite the condition's rarity in clinical practice. Furthermore, the successful validation of synthetic ECGs could address the critical challenge of data scarcity in developing AI diagnostic tools for BrS, allowing the creation of more robust and accurate algorithms through augmented training datasets.

Methods

Synthetic ECGs were generated using a Generative Adversarial Network (GAN), namely Pulse2Pulse GAN, trained on 12-lead ECG recordings from 21 Brugada patients. The detailed network architecture and implementation specifications are described in the referenced work.[15] The real ECG recordings used for training had a sampling frequency of 1 kHz, amplitude resolution of 1 µV, and each recording lasted 5 min. Data Institutional and Ethics Committee approval was obtained (BASEC 2019-00754/CE 3476). The diagnosis of Brugada syndrome was established based on the current guidelines and the Shanghai Score System criteria.[2] The synthetic ECGs were generated to match these same technical specifications, ensuring signal quality and characteristics comparable to the real recordings.

For the clinical evaluation, seven cardiologist experts in BrS (E.R.B., P.B., J.P., G.S.B., L.C., G.C., and B.B.) with >15 years of experience in the field were recruited. A dataset of 2244 ECG instances was created (50% synthetic ECGs, 50% ECGs from real BrS patients). This ratio was not disclosed to the evaluating physicians. All ECGs were presented in a standardized format using a proprietary web API with identical display settings (standard paper

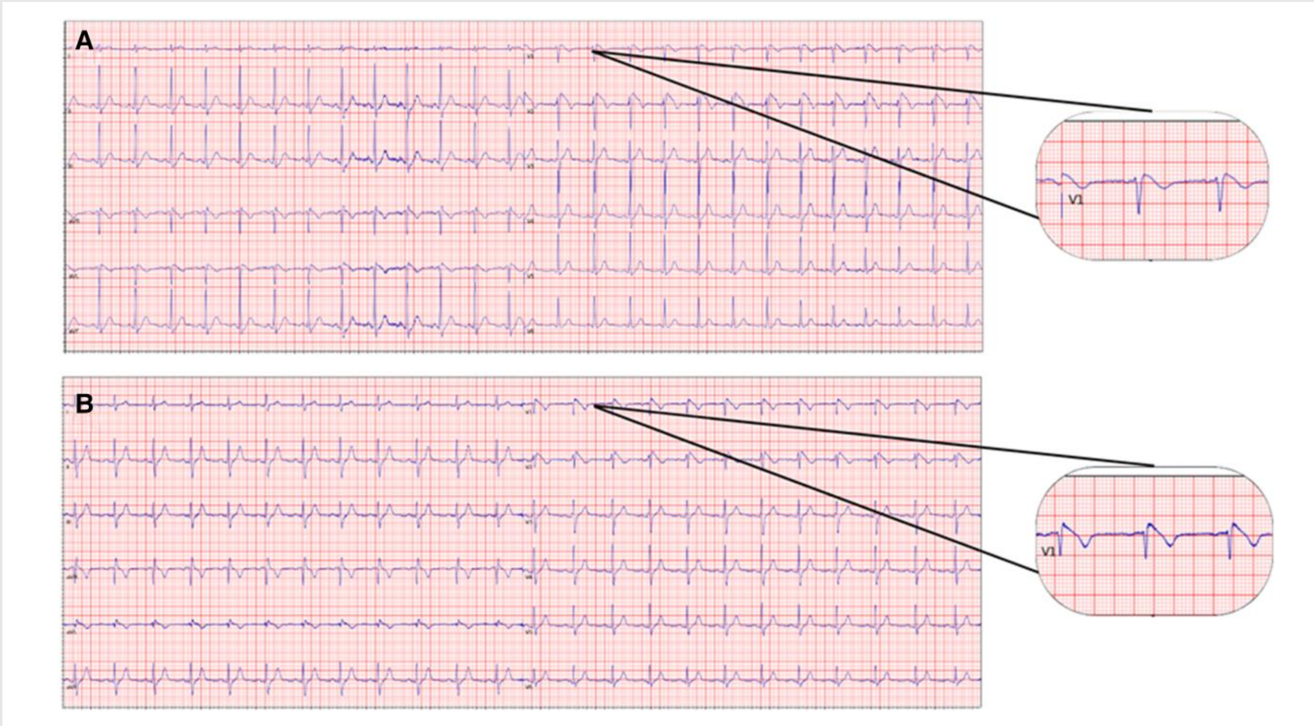


Figure 1 Comparison of *A*) real ECG and *B*) synthetic ECG from Brugada dataset.

speed of 25 mm/s, amplitude scaling of 10 mm/mV) and viewing conditions for both real and synthetic ECGs. The ECGs were randomized and displayed without any identifying markers or metadata that could indicate their origin.

The cardiologists were instructed to classify each recording either as 'real' or 'synthetic', without any additional guidance or clinical context. To investigate the repeatability of their assessments, cardiologists performed this classification task twice, with at least 2 h interval between evaluations. This repeated, blinded evaluation aimed to test both the distinguishability of synthetic ECGs and the consistency of the clinicians' judgements.

Figure 1 represents an example of real and synthetic Brugada type I ECGs. As shown, synthetic ECG highly resembled the typical features of Brugada type I pattern. Overall characteristics, like rhythmicity, ECG waveforms and cardiac axis are highly realistic in synthetic sample.

Statistical analysis and evaluation metrics

Performance metrics were calculated considering real ECGs as the positive class and synthetic ECGs as the negative class. We computed sensitivity (true positive rate, proportion of correctly identified real ECGs), specificity (true negative rate, proportion of correctly identified synthetic ECGs), precision (positive predictive value, proportion of true real ECGs among those classified as real), accuracy (proportion of correct classifications overall), and F1-score (harmonic mean of precision and sensitivity). Wilcoxon signed-rank test assessed differences in performance metrics between evaluation rounds, with statistical significance set at $\alpha = 0.05$. Inter-rater reliability between rounds was evaluated using Cohen's Kappa coefficient, which measures agreement beyond chance, with values interpreted as: < 0 poor, 0.01–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, and 0.81–1.00 almost perfect agreement. All statistical analyses were performed using Python's Scipy and Sklearn libraries.

Results

Experienced cardiologists' evaluation of synthetic compared with real ECG instances revealed high challenges in distinguishing between the two types of data. In the first evaluation round, the cardiologists achieved an overall mean accuracy of 40%, with similar sensitivity (58%, indicating the ability to correctly identify real ECGs) but lower specificity (22%). The second round documented a slight improvement in overall accuracy (42%), specificity (26%), and precision (43%), with minimal change in sensitivity (58%) and a small increase in the F1-score (50%). Overall, the ECG evaluation showed consistently higher sensitivity than specificity across both rounds (58% vs. 22% in Round 1, 58% vs. 26% in Round 2). This pattern suggests a notable tendency to classify ECGs as real, even when they are synthetic, implying that synthetic ECGs closely resemble real ones. The high sensitivity, paired with lower specificity, reflects the realistic features of the synthetic data generated.

Table 1 Performance metrics of cardiologists in distinguishing synthetic against real ECGs across two evaluation rounds

Metric	Round 1	Round 2	*P*-value
Accuracy	40 ± 18	42 ± 13	0.71
Sensitivity	58 ± 25	58 ± 16	1.0
Specificity	22 ± 12	26 ± 13	0.62
Precision	41 ± 16	43 ± 10	0.62
F1-score	48 ± 19	50 ± 12	0.81

Results are presented as mean ± standard deviation obtained across scorers.

Wilcoxon signed-rank test comparing performance metrics between the two rounds showed no statistically significant differences (all *P*-values > 0.6) *Table 1*. Intra-rater reliability varied widely, with Cohen's Kappa values ranging from 0.80 (indicating substantial agreement) for one cardiologist to −0.12 (indicating disagreement) for another, suggesting notable inconsistency in individual assessments. These findings highlight the difficulty in reliably distinguishing synthetic ECGs from real ones and the variability in cardiologist performance over repeated evaluations.

Discussion

Our findings reveal a concerning difficulty of experienced cardiologists to consistently distinguish between real and synthetic Brugada ECGs, as denoted by low accuracy of 40% and 42% in Round 1 and Round 2, respectively. This raises significant implications for both clinical practice, research methodology and may overcome the machine learning issue related to data scarcity in patients with rare cardiac diseases. Different previous studies have reported similar results.[14] Alcaraz *et al.*[16] exploited a diffusion-based conditional ECG generator, and a single expert cardiologist validated a small subset of synthetic data. Result showed low accuracy (50%), specificity (22.75%) and precision (50%) while a modest sensitivity was reached (77%).[16] Similarly, Cao *et al.*[17] validated a generative network through the aid of an expert technician, who obtained an overall accuracy in distinguishing real and generated ECGs of 63% (normal sinus rhythm samples) and 75% (atrial fibrillation samples).[17] The ease with which clinicians are misled by synthetic ECGs underscores the sophisticated nature of current ECG generation techniques, but also highlights potential risks in relying on artificial data without proper disclosure.

In the current work, the observed inconsistency in repeated assessments further emphasizes the challenge synthetic ECGs pose to clinical judgement, with Cohen's Kappa values ranging from 0.80 to −0.12. This uncertainty could have far-reaching consequences in diagnostic accuracy and treatment decisions if synthetic ECGs are inadvertently integrated into clinical workflows or training materials without clear identification.

However, it is crucial to recognize that synthetic ECGs, when properly managed and transparently used, offer substantial benefits to cardiac research and education. They can augment limited real-world datasets, enabling more comprehensive training for medical professionals and facilitating the development of improved diagnostic algorithms. Synthetic data can also address privacy concerns associated with using real patient ECGs, potentially accelerating research, and innovation in cardiac care. The key lies in striking a balance between leveraging the advantages of synthetic ECGs and mitigating the risks they pose.

The upcoming challenge is to develop robust protocols for the generation, use, and disclosure of synthetic ECGs. This approach would allow the field to harness the full potential of this technology while maintaining the integrity of clinical practice and research methodology. Researchers must explicitly declare the use of synthetic data in all publications and presentations. Clear guidelines are needed for the integration of synthetic ECGs in research datasets to maintain data integrity. Studies involving synthetic ECGs should include this information in participant consent processes. The development of standardized protocols for the use and disclosure of synthetic ECGs in both research and clinical training contexts is advised.

Moving forward, several key areas require attention. There is a need to improve synthetic ECG generation techniques to create more distinguishable artificial ECGs, potentially incorporating deliberate artefacts that signal their artificial nature. The potential for AI-driven tools to assist in distinguishing between real and synthetic ECGs should be investigated. Lastly, collaboration with regulatory bodies is crucial to establish industry-wide standards for the use and disclosure of synthetic medical data. These steps are crucial to harness the benefits of synthetic ECGs in research and education while mitigating risks to clinical practice and scientific integrity.

The scientific community must work collaboratively to establish clear guidelines for the generation, use, and disclosure of synthetic ECGs. By doing so, we can harness the potential of this technology to enhance learning, expand research capabilities, and ultimately improve patient care, while safeguarding the integrity of clinical practice and scientific inquiry.

Our study presents some limitations that warrant consideration. While our findings demonstrate the difficulty in distinguishing between real and synthetic ECGs, these results are specifically obtained for BrS ECGs generated by the exploited generative AI-model (Pulse2Pulse GAN). The observed indistinguishability may not generalize to other cardiac conditions or different ECG generation techniques. Further research is needed to validate these findings across different cardiac pathologies and generative AI-tools.

Conclusions

Brugada expert cardiologists cannot distinguish synthetic Brugada ECGs from real BrS patients' ECG.

Author contributions

B. Zanchi (PhD): Conceptualization, Methodology, Formal analysis, Visualization, and Writing – Original draft; G. Monachino (PhD): Methodology, Formal analysis, and Review & editing; F.D. Faraci (PhD): Project administration and Review & editing; M. Metaldi (MD, PhD): Formal analysis, Software, and Review & editing; P. Brugada (MD, PhD), G. Sarquella-Brugada (MD, PhD), E.R. Behr (MA, MBBS, MD), J. Brugada (MD, PhD), L. Crotti (MD, PhD), & B. Belhassen (MD, PhD): Investigation and Review & editing; G. Conte (MD, PhD): Data curation, Funding acquisition, Investigation, Resources, Project administration, and Writing – Review & editing.

Funding

This study was supported by a research grant of the Swiss National Science Foundation (PZ00P3_180055).

Conflict of interest: none declared.

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

References

1. Crotti L, Brugada P, Calkins H, Chevalier P, Conte G, Finocchiaro G, *et al.* From gene-discovery to gene-tailored clinical management: 25 years of research in channelopathies and cardiomyopathies. *Europace* 2023;**25**:euad180.
2. Zeppenfeld K, Tfelt-Hansen J, De Riva M, Winkel BG, Behr ER, Blom NA, *et al.* 2022 ESC guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death. *Eur Heart J* 2022;**43**:3997–4126.
3. Liao S, Bokhari M, Chakraborty P, Poudel D, Cibils M, Thomas J, *et al.* Use of wearable technology and deep learning to improve the diagnosis of Brugada syndrome. *JACC Clin Electrophysiol* 2022;**8**:1010–1020.
4. Liu CM, Liu CL, Hu KW, Lin YH, Lin KH, Wang JS, *et al.* A deep learning-enabled electrocardiogram model for the identification of a rare inherited arrhythmia: Brugada syndrome. *Can J Cardiol* 2022;**38**:152–159.
5. Melo L, Ciconte G, Christy A, Faraci FD, Conte G, Saeed Y, *et al.* Deep learning unmasks the ECG signature of Brugada syndrome. *PNAS Nexus* 2023;**2**:pgad327.
6. Micheli A, Natali M, Pedrelli L, Dimitri GM, Vozzi F, Vozzi G. Analysis and interpretation of ECG time series through convolutional neural networks in Brugada syndrome diagnosis. In: International Conference on Artificial Neural Networks, Crete, Greece. Springer; 2023. p26–36.
7. Zanchi B, Faraci FD, Gharaviri A, Bergonti M, Monga T, Auricchio A, *et al.* Identification of Brugada syndrome based on P-wave features: an artificial intelligence-based approach. *Europace* 2023;**25**:euad334.

8. Leong CJ, Sharma S, Seth J, Rabkin SW. Artificial intelligence streamlines diagnosis and assessment of prognosis in Brugada syndrome: a systematic review and meta-analysis. *Connect Health Telemed* 2024; **3**.
9. Vozzi F, Pedrelli L, Dimitri GM, Natali M, Micheli A, Vozzi G. Echo state networks for the recognition of type 1 Brugada syndrome from conventional 12-LEAD ECG. *Heliyon* 2024;**10**:e25404.
10. Calburean PA, Pannone L, Monaco C, Ciconte G, Monasky MM, Micaglio E, *et al.* Predicting and recognizing drug-induced type I Brugada pattern using ECG-based deep learning. *J Am Heart Assoc* 2024;**13**:e033148.
11. Yoo H, Moon J, Kim JH, Joo HJ. Design and technical validation to generate a synthetic 12-lead electrocardiogram dataset to promote artificial intelligence research. *Health Inf Sci Syst* 2023; **11**:41.
12. Winkler MA, Herrmann Von, Brooks PF, Hoopes MA, Attili CW, Sorrell A, *et al.* Synthetic ECG-gated cardiac computed tomography for in vivo imaging of the temporary total artificial heart. *Circulation* 2013; **127**:e4–e5.
13. Monachino G, Zanchi B, Fiorillo L, Conte G, Auricchio A, Tzovara A, *et al.* Deep generative models: the winning key for large and easily accessible ECG datasets? . *Comput Biol Med* 2023;**167**:107655.
14. Zanchi B, Monachino G, Fiorillo L, Conte G, Auricchio A, Tzovara A, *et al.* Synthetic ECG signals generation: a scoping review. *Comput Biol Med* 2025;**184**: 109453.
15. Thambawita V, Isaksen JL, Hicks SA, Ghouse J, Ahlberg G, Linneberg A, *et al.* DeepFake electrocardiograms using generative adversarial networks are the beginning of the end for privacy issues in medicine. *Sci Rep* 2021;**11**:21896.
16. Alcaraz JML, Strodthoff N. Diffusion-based conditional ECG generation with structured state space models. *Comput Biol Med* 2023;**163**:107115.
17. Cao F, Budhota A, Chen H, Rajput KS. Feature matching based ECG generative network for arrhythmia event augmentation. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, Canada. IEEE; 2020. p296–299.