



## LJMU Research Online

Essa, H, Ortega-Martorell, S, Olier, I and Lip, GYH

**Machine learning to identify phenotypic clusters of patients with atrial fibrillation**

<http://researchonline.ljmu.ac.uk/id/eprint/25094/>

### Article

**Citation** (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

**Essa, H, Ortega-Martorell, S, Olier, I and Lip, GYH (2024) Machine learning to identify phenotypic clusters of patients with atrial fibrillation. Heart Rhythm O2, 6 (2). pp. 139-141. ISSN 2666-5018**

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact [researchonline@ljmu.ac.uk](mailto:researchonline@ljmu.ac.uk)

<http://researchonline.ljmu.ac.uk/>

# Machine learning to identify phenotypic clusters of patients with atrial fibrillation



Hani Essa, MBChB,<sup>1</sup> Sandra Ortega-Martorell, PhD,<sup>1,2</sup> Ivan Olier, PhD,<sup>1,2</sup>  
Gregory Y.H. Lip, MD<sup>1,3,4</sup>

From the <sup>1</sup>Liverpool Centre for Cardiovascular Science at University of Liverpool, Liverpool Heart and Chest Hospital, Liverpool John Moores University, Liverpool, United Kingdom, <sup>2</sup>Data Science Research Centre, Liverpool John Moores University, Liverpool, United Kingdom, <sup>3</sup>Danish Center for Health Services Research, Department of Clinical Medicine, Aalborg University, Aalborg, Denmark, and <sup>4</sup>Department of Cardiology, Lipidology and Internal Medicine, Medical University of Bialystok, Bialystok, Poland.

Atrial fibrillation (AF) is the most common cardiac arrhythmia worldwide and is associated with significant morbidity and mortality, including stroke, heart failure, dementia, and hospitalizations.<sup>1</sup> As a result, better efforts to identify patients at greatest risk, who would benefit most from appropriate management, are needed.

Oral anticoagulation can be used to ameliorate the risk of stroke in AF, but the management of AF is more than just oral anticoagulation, given the recognized residual risks of major adverse events despite anticoagulation.<sup>2</sup> Indeed, AF is not a homogeneous single diagnosis, and over recent years, phenotypes of “clinical complexity” associated with AF have been identified, with implications for prognosis and management.<sup>3</sup>

The current management of AF has moved toward a more holistic or integrated care approach, initially proposed as the ABC (Atrial fibrillation Better Care) pathway.<sup>4</sup> The ABC pathway is supported by trial and real-world evidence,<sup>5</sup> and variants of the “ABC” acronym have been used in U.S. guidelines (ie, SOS [Stroke, Other Comorbidities, Rate or Rhythm control])<sup>6</sup> and 2024 European guidelines (as CARE [ie, Comorbidities, Avoid stroke, Rate or rhythm control, Evaluation]).<sup>7</sup>

Phenotypic clusters of AF patients identified by hierarchical cluster analysis show improved outcomes with ABC pathway adherence but to a varying degree depending on their phenotype.<sup>3</sup> Adherence to the ABC pathway has been associated with a significant reduction in all-cause mortality, cardiovascular mortality, stroke, and bleeding.<sup>8</sup>

Turning to stroke risk, the more common and well-validated risk factors have been utilized to formulate stroke risk stratification schemes, of which the one most used is the

CHA<sub>2</sub>DS<sub>2</sub>-VASc (congestive heart failure, hypertension, age  $\geq 75$  years, diabetes mellitus, prior stroke or transient ischemic attack or thromboembolism, vascular disease, age 65–74 years, sex category) score.<sup>9</sup> While clinically useful to identify patients who may benefit from anticoagulation, this is an oversimplification of a much more complex and dynamic scenario, and hence only demonstrates a modest predictive performance of stroke risk. There is a clinically apparent need for better risk-stratification strategies to identify patients who may benefit from anticoagulation.

Beyond hierarchical cluster analysis, other approaches such as latent class analysis have been used to phenotype patients with AF.<sup>10</sup> Nevertheless, artificial intelligence (AI) and machine learning (ML), a subset of AI, signal the emergence of tools that can help us leverage large data sets to identify clinically significant patterns that may not be easily identified by conventional methods, and this is demonstrated in dramatic growth in the numbers of publications in this field over the last few years.<sup>11</sup>

In this issue of *Heart Rhythm O<sup>2</sup>*, Hsu and colleagues<sup>12</sup> used a statistical approach to identify distinct prognostic phenotypic clusters in a Taiwanese population of 5002 patients with AF. In this analysis, the authors perform an unsupervised hierarchical cluster analysis based on the components of the CHA<sub>2</sub>DS<sub>2</sub>-VASc score, identifying 4 distinct clusters: cluster I included 1918 diabetic patients with heart failure with preserved ejection fraction, and chronic kidney disease; cluster II comprised 1006 older patients with low body mass index and pulmonary hypertension; cluster III consisted of 1731 patients with metabolic syndrome and atherosclerotic disease; and cluster IV included 347 patients with left heart dysfunction, including reduced ejection fraction.

The main outcomes measured across all clusters were the risk of ischemic stroke, heart failure hospitalization, cardiovascular death, and all-cause mortality. First, Hsu and

---

**Address reprint requests and correspondence:** Dr Gregory Y.H. Lip, Liverpool Centre for Cardiovascular Science, William Henry Duncan Building, 6 West Derby Street, Liverpool L7 8TX, UK. E-mail address: [gregory.lip@liverpool.ac.uk](mailto:gregory.lip@liverpool.ac.uk); Twitter: @LiverpoolCCS.

colleagues found significant differences in the risk of ischemic stroke independent of CHA<sub>2</sub>DS<sub>2</sub>-VASc score between clusters, with cluster IV demonstrating the lowest risk. Second, cluster II was associated with the highest risk of heart failure hospitalization, cardiac death, and all-cause mortality. Finally, the data-driven algorithm identified heterogeneous risk profiles across different clusters, each associated with a varying risk of cardiovascular events. The study by Hsu and colleagues identified distinct phenotypes that demonstrated a differential risk of stroke, independent of their CHA<sub>2</sub>DS<sub>2</sub>-VASc score. The authors further benefitted from a large sample size and a long duration of follow-up. Furthermore, these results were externally validated on a separate dataset, lending credibility and demonstrating reproducibility of the findings.

Nonetheless, several limitations must be considered when interpreting these results. First, this is a retrospective study in a predominantly homogeneous East Asian population and is therefore open to confounding factors that may explain these findings. Second, as a study dependent on data extraction from an administrative database (without a review of individual patient charts), there is a significant risk that data are subject to coding errors, which can alter results. Last, even though their work is supported by external validation on a Taiwanese dataset, its applicability to non-Asian populations is uncertain, especially given the reported racial differences in AF-related outcomes such as stroke and bleeding.<sup>13,14</sup>

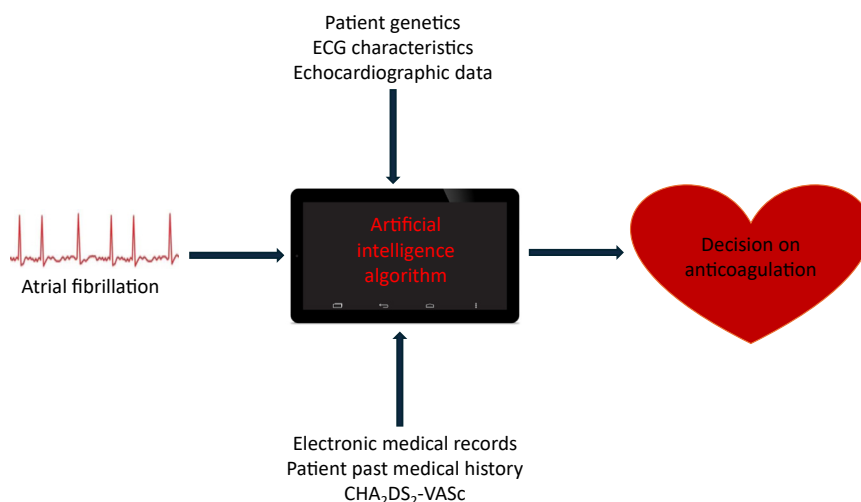
The work conducted by Hsu and colleagues is timely and contributes to the expanding role of data-driven approaches in the management of AF.<sup>11</sup> Moving forward, phenotyping clinically complex AF patients can deploy more ML sophisticated approaches, such as generative topographic mapping,

as recently published by our group.<sup>15</sup> AI/ML has the potential to be able to provide a continuous “real-time” assessment of individual risk in AF outperforming traditional stroke risk stratification schemes.<sup>16</sup> This is augmented by the growth in methodologies such as digital twins, currently applied in extensive research programs to improve diagnosis, risk prediction, peristroke management, and poststroke rehabilitation.<sup>17,18</sup>

However, risk factors and comorbidities in AF patients are not static, but rather are dynamic in nature, and the arrhythmia per se is also dynamic, changing in patterns over time.<sup>19,20</sup> Furthermore, we have adherence to the ABC pathway that is also dynamic, and adherence/nonadherence over follow-up can impact outcomes.<sup>21</sup>

Novel ML approaches could help identify phenotypic clusters of AF patients, who have a high risk of ischemic stroke, despite being deemed low risk by traditional risk scores such as CHA<sub>2</sub>DS<sub>2</sub>-VASc and randomizing these patients to ABC pathway-based management vs conventional treatment. The ability to identify these high-risk individuals who would otherwise be missed could allow a tailored approach to decision making and improve overall patient care. Ultimately, we may utilize AI/ML to create an algorithm that incorporates conventional patient data used in normal risk stratification schemes combined with nonconventional and dynamic data for the identification of which patients may benefit from anticoagulation (Figure 1).

In conclusion, the growth in AI/ML yields promising results for the identification of high-risk patients who may otherwise be missed via conventional stroke stratification schemes. Integrating AI/ML into the diagnostic and treatment processes for AF has the potential to mitigate current



**Figure 1** The potential role of artificial intelligence and machine learning in decision making for anticoagulation. CHA<sub>2</sub>DS<sub>2</sub>-VASc = congestive heart failure, hypertension, age  $\geq 75$  years, diabetes mellitus, prior stroke or transient ischemic attack or thromboembolism, vascular disease, age 65–74 years, sex category; ECG = electrocardiography.

limitations and optimize care. The “rise of the machines” is clearly evident in healthcare.

**Funding Sources:** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Disclosures:** Hani Essa discloses no conflicts. Sandra Ortega-Martorell is the principal investigator of the TARGET project on health virtual twins for personalized management of atrial fibrillation and stroke (grant agreement no. 101136244) and senior investigator in the ARISTOTELES project on artificial intelligence for the management of chronic long-term conditions (grant agreement no. 101080189), both funded by the European Union’s Horizon Europe Research and Innovation programme; and has served on the board of the ART (Ageing Research Translation) of the Healthy Ageing Network funded by the Biotechnology and Biological Sciences Research Council. Ivan Olier is the methodological lead of the TARGET project on health virtual twins for personalized management of atrial fibrillation and stroke (grant agreement no. 101136244) and partner lead in the ARISTOTELES project on artificial intelligence for the management of chronic long-term conditions (grant agreement no. 101080189), both funded by the European Union’s Horizon Europe Research and Innovation programme. Gregory Y.H. Lip is a National Institute for Health and Care Research Senior Investigator; has served as a consultant and speaker for BMS/Pfizer, Boehringer Ingelheim, Daiichi Sankyo, and Anthos (no fees are received personally); is co-principal investigator/lead of the AFFIRMO project on multimorbidity in AF (grant agreement no. 899871), TARGET project on digital twins for personalized management of atrial fibrillation and stroke (grant agreement no. 101136244), and ARISTOTELES project on artificial intelligence for management of chronic long-term conditions (grant agreement no. 101080189), which are all funded by the European Union’s Horizon Europe Research and Innovation programme.

**Authorship:** All authors attest they meet the current ICMJE criteria for authorship.

## References

- Burdett P, Lip GYH. Atrial fibrillation in the UK: predicting costs of an emerging epidemic recognizing and forecasting the cost drivers of atrial fibrillation-related costs. *Eur Heart J Qual Care Clin Outcomes* 2022;8:187–194.
- Liu Y, Chen Y, Olier I, et al; GLORIA-AF Investigators. Residual risk prediction in anticoagulated patients with atrial fibrillation using machine learning: a report from the GLORIA-AF registry phase II/III. *Eur J Clin Invest*. Published online December 11, 2024; <https://doi.org/10.1111/eci.14371>.
- Krittayaphong R, Treewaree S, Kaewkumdee P, Lip GYH. Clinical phenotype classification to predict risk and optimize the management of patients with atrial fibrillation using the Atrial Fibrillation Better Care (ABC) pathway: a report from the COOL-AF registry. *QJM* 2024;117:16–23.
- Lip GYH. The ABC pathway: an integrated approach to improve AF management. *Nat Rev Cardiol* 2017;14:627–628.
- Potpara T, Romiti GF, Sohns C. The 2024 European Society of Cardiology guidelines for diagnosis and management of atrial fibrillation: a viewpoint from a practicing clinician’s perspective. *Thromb Haemost* 2024;124:1087–1094.
- Joglar JA, Chung MK, Armbruster A, et al. 2023 ACC/AHA/ACCP/HRS guideline for the diagnosis and management of atrial fibrillation: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation* 2024;149:e1–e156.
- Van Gelder IC, Rienstra M, Bunting KV, et al. 2024 ESC guidelines for the management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS). *Eur Heart J* 2024; 45:3314–3414.
- Romiti GF, Pastori D, Rivera-Caravaca JM, et al. Adherence to the ‘Atrial Fibrillation Better Care’ Pathway in Patients with atrial fibrillation: impact on clinical outcomes—a systematic review and meta-analysis of 285,000 patients. *Thromb Haemost* 2022;122:406–414.
- Lip GY, Nieuwlaar R, Pisters R, Lane DA, Crijns HJGM. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. *Chest* 2010;137:263–272.
- Romiti GF, Corica B, Mei DA, et al. Patterns of comorbidities in patients with atrial fibrillation and impact on management and long-term prognosis: an analysis from the Prospective Global GLORIA-AF Registry. *BMC Med* 2024; 22:151.
- Olier I, Ortega-Martorell S, Pieroni M, Lip GYH. How machine learning is impacting research in atrial fibrillation: implications for risk prediction and future management. *Cardiovasc Res* 2021;117:1700–1717.
- Hsu J-C, Yang Y-Y, Chuang S-L, Lin L-Y. Phenotypes of atrial fibrillation in a Taiwanese longitudinal cohort: Insights from an Asian perspective. *Heart Rhythm O2* 2025;6:129–138.
- Kang DS, Yang P-S, Kim D, et al. Racial differences in bleeding risk: an ecological epidemiological study comparing Korea and United Kingdom subjects. *Thromb Haemost* 2024;124:842–851.
- Kang DS, Yang P-S, Kim D, et al. Racial differences in ischemic and hemorrhagic stroke: an ecological epidemiological study. *Thromb Haemost* 2024;124:883–892.
- Bellfield RAA, Olier I, Lotto R, et al. AI-based derivation of atrial fibrillation phenotypes in the general and critical care populations. *EBioMedicine* 2024; 107:105280.
- Lip GYH, Genaidy A, Tran G, Marroquin P, Estes C, Sloop S. Improving stroke risk reduction in the general population: a comparative assessment of common clinical rules, a new multimorbid index, and machine-learning-based algorithms. *Thromb Haemost* 2022;122:142–150.
- Ortega-Martorell S, Olier I, Ohlsson M, Lip GYH; TARGET Consortium. TARGET: a major european project aiming to advance the personalized management of atrial fibrillation-related stroke via the development of health virtual twins technology and artificial intelligence. *Thromb Haemost*. Published online November 7, 2024; <https://doi.org/10.1055/a-2438-5671>.
- Ortega-Martorell S, Olier I, Lip GYH; TARGET consortium. A European network to develop virtual twin technology for personalized stroke management in atrial fibrillation: the TARGET consortium. *Eur Heart J*. Published online November 20, 2024; <https://doi.org/10.1093/eurheartj/ehae673>.
- Krittayaphong R, Winijkul A, Methavigul K, Chichareon P, Lip GYH. Clinical outcomes of patients with atrial fibrillation in relation to multimorbidity status changes over time and the impact of ABC pathway compliance: a nationwide cohort study. *J Thromb Thrombolysis*. Published online July 9, 2024; <https://doi.org/10.1007/s11239-024-03007-9>.
- Imberti JF, Bonini N, Tosetti A, et al. Atrial high-rate episodes detected by cardiac implantable electronic devices: dynamic changes in episodes and predictors of incident atrial fibrillation. *Biology (Basel)* 2022;11:443.
- Krittayaphong R, Chichareon P, Methavigul K, Treewaree S, Lip GYH. Relation of changes in ABC pathway compliance status to clinical outcomes in patients with atrial fibrillation: a report from the COOL-AF registry. *Eur Heart J Qual Care Clin Outcomes*. Published online May 24, 2024; <https://doi.org/10.1093/ehjqcco/qcae039>.