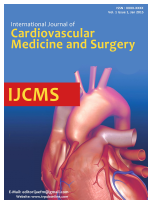
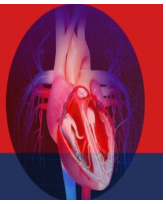
https://doi.org/10.62587/AFJCCM.2.1.2026.1-9 ISSN: 3136-9294



African Journal of Cardiology and Cardiovascular Medicine

Journal homepage: https://www.afjccm.com/



Research Article **Open Access**

Multi-omics integration for risk stratification of post-operative atrial fibrillation after cardiac surgery

Suriya K. Wijaya[1*]

[1]Nusantara Institute of Cardiac Sciences, 18 Garuda Medical Road, Jakarta, Indonesia.

Article Info

Volume 2, Issue 1, January 2026
Received : 11 August 2025
Accepted : 21 November 2025
Published : 25 January 2026
doi: 10.62587/AFJCCM.2.1.2026.1-9

Abstract

Aim: The aim of this study was to evaluate if the pre-operative risk stratification of post-operative atrial fibrillation can be improved by combining multiple omics layers. **Methods:** Patients with cardiac surgery who had not had a previous diagnosis of Atrial Fibrillation (AF) underwent pre-operative multi-omics profiling by four modalities, clinical, genomic, plasma proteomic and metabolomic. Cross validation was used to develop the per-layer risk models and the stacking integration model was used to combine the models. Each single-omics model and the clinical score were compared to the integrated model using the area under receiver operating characteristic curve, and the common directions of variation extracted through multi-omics factor analysis. **Results:** In total, 760 patients had post-operative AF in 35%. It was discriminated by the integrated multi-omics model with area under the curve of 0.82 with highly significant differences compared with best single-omics model and clinical score ($p<0.001$). This shared latent axis, which mainly consisted of inflammatory proteins and metabolic markers, discerned the patients developing atrial fibrillation, factor analysis found. Protein biomarkers and metabolite biomarkers were the top features of the signature. **Conclusion:** The combined clinical, genome, proteome and metabolome data performed better than any single layer and better than the clinical score alone in predicting post-operative atrial fibrillation, and uncovered an inflammatory and metabolic signature by the omics, demonstrating the potential of multi-omic integration in improving the clinical risk stratification of perioperatively.

Keywords: *Postoperative atrial fibrillation, Cardiac surgery, Multi-omics, Proteomics, Metabolomics, Risk prediction, Data integration*

© 2026 Suriya K. Wijaya. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

1. Introduction

As the most frequently encountered complication after cardiac surgery, atrial fibrillation is a common occurrence. Predicting and preventing it are important, as it occurs in approximately one in three patients in the first few

* *Corresponding author: Suriya K. Wijaya, Nusantara Institute of Cardiac Sciences, 18 Garuda Medical Road, Jakarta, Indonesia.*

3136-9294/© 2026 Suriya K. Wijaya. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

days following surgery, although it is usually temporary, it is not harmless, it is associated with stroke, longer stay and poorer survival (Dobrev *et al.*, 2019; Greenberg *et al.*, 2017). A bit of a problem is it's not easy to predict. The clinical scores that have been engineered to predict it (including according to with atrial size and type of surgery) only account for part of the risk and discern poorly, such that, clinicians are still unable to confidently predict which patient will fibrillate and which patient to preserve (Mariscalco *et al.*, 2014; Hindricks *et al.*, 2021).

The effort of predicting the burden is justified because of its burden. In addition to the acute haemodynamic disturbance, an episode of post-operative atrial fibrillation is known to prolong intensive care stay, hospital stay, the cost of admission and, most importantly, it is linked to an increased risk of stroke and non-perioperative-related deaths as well as a poor trajectory, some of which may result from the arrhythmia. Sometimes prophylaxis does exist such as Beta blockade, amiodarone and others, but because of the risks involved in prophylaxis it should not sensibly be given to everyone, so the use of it is going to rely on being targeted to the patients who are true fibillers – which is exactly where prediction is missing. A more suitable estimation of risk is therefore not an end in itself but rather is the precondition needed in order to use prevention well.

The difficulty of a clinical score could stem from the fact that the disorder has more than one cause. Post-operative atrial fibrillation is a condition in which a vulnerable atrium is coupled to the noxious exposure of surgery: the genome code sets an exposure of the atrium to the lifelong risk to fibrillation, the inflammation introduces the noxious stimulus to the vulnerable atrium, and the balance of metabolism of the myocardium contributes to the stress of surgery (Roselli *et al.*, 2018). A marker of these, an inflammatory protein, a metabolite, a risk allele, reveals only their respective piece in a variety of languages.

The concept of knowledge clustering through multiple layers of molecular understanding is based on a simple premise: none of these layers describes the whole patient. The genome is a fixed predisposition not the state of inflammation; the proteome is the report of the state of inflammation, but not of the genome, the metabolome is a report of the current metabolic states, but not of the structural site of its action. They are each partial views, and a disease that comes from their interplay will only be discernible in one of these views (Argelaguet *et al.*, 2018; Picard *et al.*, 2021). There is a promise for the multi-omics integration that is highest at the intersection of multiple biological systems (and not single systems) for the disorder being studied, such as the post-operative atrial fibrillation studies mentioned above.

This is the problem that multi-omics integration is designed for. By measuring several molecular layers in the same patient, the genome, the proteome, the metabolome, and combining them with a method that finds their shared and complementary signals, one can in principle assemble a fuller picture of risk than any layer gives alone (Singh *et al.*, 2019). The approach has matured in cancer and in other complex diseases, with established methods for the unsupervised discovery of shared axes of variation and for the supervised integration of layers into a predictor, and it is now being brought to the perioperative setting, where the pre-operative window offers a natural moment to profile a patient across omics before the operation that will test the atrium (Reel *et al.*, 2021).

This study applied multi-omics integration to the pre-operative risk stratification of post-operative atrial fibrillation. In a cohort of cardiac surgery patients profiled before operation across clinical, genomic, proteomic and metabolomic layers, it built a risk model from each layer, integrated them, and tested whether the integrated model improved on the single layers and on the clinical score, then characterised the shared biology the integration revealed.

2. Materials and methods

2.1. Study design and population

This was a retrospective multi-omics study in a cohort of adults undergoing cardiac surgery, coronary, valve, or combined, who had no history of atrial fibrillation and who were in sinus rhythm before operation. The study was reported in accordance with the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis guidance (Collins *et al.*, 2015). A total of 760 patients were studied.

2.2. Pre-operative multi-omics profiling

Each patient was profiled before operation across four layers. The clinical layer comprised the variables of established risk, including age, the left atrial diameter, the ejection fraction and the type of operation. The genomic layer was summarised by a polygenic risk score for atrial fibrillation, computed from genome-wide genotyping, imputed to a reference panel, and weighted by the effect sizes of a large atrial fibrillation genome-wide association study (Weng *et al.*, 2018). The proteomic layer was measured in pre-operative plasma by the proximity extension assay, giving 42 cardiovascular and inflammatory proteins on a normalised log2 scale (Assarsson *et al.*, 2014). The metabolomic layer was measured in the same sample by mass spectrometry, giving 34 metabolites spanning acylcarnitines, ketone bodies, amino acids and microbial co-metabolites (Wishart, 2019).

2.3. Definition of post-operative atrial fibrillation

The outcome was new-onset atrial fibrillation in the post-operative period, detected on continuous telemetry or the twelve-lead electrocardiogram and lasting long enough to be recorded clinically, in a patient with no prior history of the arrhythmia.

2.4. Per-omics predictive modelling

Within each omics layer a model was built to predict post-operative atrial fibrillation. The proteomic and metabolomic layers, being high-dimensional, were modelled by gradient-boosted decision trees, and the clinical and genomic layers by logistic regression, each feature being standardised beforehand (Chen and Guestrin, 2016). Every per-layer model was trained and evaluated by five-fold cross-validation, and the out-of-fold predicted probabilities were retained as that layer's risk score, so that no patient's score was informed by a model trained on that patient.

2.5. Multi-omics integration

The four layers were integrated by stacked generalisation, in which the out-of-fold risk scores of the per-layer models become the inputs to a second-level model that learns how best to combine them (Wolpert, 1992). The second-level model was a logistic regression, itself fitted and evaluated by five-fold cross-validation on the layer scores, yielding an integrated risk score for each patient that draws on all four layers while remaining protected against over-fitting by the nested cross-validation. This late-integration strategy allows each layer to be modelled in the way that suits it before the layers are combined.

2.6. Multi-omics factor analysis

To characterise the shared structure across the layers, the standardised features of all four were analysed together by factor analysis, in the spirit of multi-omics factor analysis, which decomposes the joint variation into a small number of latent factors. For each factor the proportion of its variance contributed by each omics layer was computed from the squared loadings, and the association of each factor with post-operative atrial fibrillation was tested, so that the factors carrying the outcome signal, and the layers driving them, could be identified.

2.7. Model evaluation and statistical analysis

The area under the ROC curve, test by bootstrap difference of area, 95 percent confidence interval, bootstrap re-sample untransformed, was used to quantify discrimination for each single-omics model and for the clinical score, and to compare the integrated model with each single-omics model and with the clinical score (DeLong *et al.*, 1988). Given the imbalanced nature of the outcome, the precision-recall curve was also calculated, and the fit of the integrated model was evaluated in each of 10 intervals defined by predicted risk.

3. Results

3.1. Study population

Of the 760 patients, 265 (35%) developed post-operative atrial fibrillation. Those who did were older (61.1 versus 55.9 years), had a larger left atrium (41.5 versus 39.4 mm) and a slightly lower ejection fraction, the expected clinical profile, as given in Table 1; the differences, though real, were modest, consistent with the limited discrimination of clinical features alone.

Table 1: Characteristics of patients with and without post-operative atrial fibrillation

Characteristic	POAF	No POAF
Number	265	495
Age, years (mean)	61.1	55.9
Left atrial diameter, mm	41.5	39.4
LV ejection fraction, %	53	56
Valve surgery, %	45	42

3.2. Discrimination and multi-omics integration

The integrated multi-omics model discriminated post-operative atrial fibrillation well, with an area under the curve of 0.82 (95% CI 0.79 to 0.85), and it exceeded every alternative, as shown in Figure 1 and given in Table 2. Among the single layers the proteomic and metabolomic models were the strongest (0.75 and 0.75), the genomic score alone was weak (0.60), and the clinical score reached only 0.66, in keeping with the known limits of clinical prediction here. The integrated model improved significantly on the best single layer (difference in area 0.06; $p<0.001$), the gain of integration being to combine the complementary signals of the layers rather than to rely on any one.

Table 2: Discrimination of post-operative atrial fibrillation by model

Model	AUROC	95% CI
Integrated multi-omics	0.82	0.79–0.85
Proteomic	0.75	0.72–0.79
Metabolomic	0.75	0.71–0.79
Clinical score	0.66	0.62–0.70
Genomic (polygenic score)	0.60	0.56–0.65

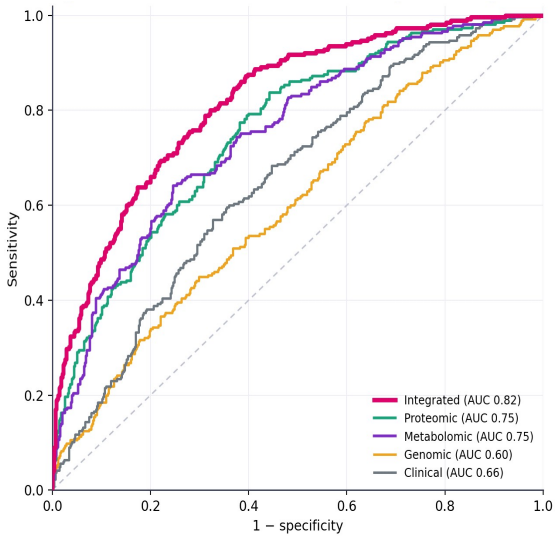


Figure 1: Discrimination of post-operative atrial fibrillation by the integrated model and the comparators

The two molecular layers, proteomic and metabolomic, discriminated almost equally and better than the clinical or genomic layers, yet their combination in the integrated model exceeded either, indicating that they carried partly distinct rather than redundant information about the outcome. Had the two simply measured the same inflammatory state, integrating them would have added little; that integration improved on each implies that the metabolomic layer captured metabolic derangement not fully reflected in the proteome, and the proteomic layer inflammation not fully reflected in the metabolome, so that each supplied what the other missed.

3.3. The integrated latent space

Four layers were combined to do the factor analysis which yielded the latent factors as shown in Table 3 with one factor correlating very well with the outcome. The patients who developed atrial fibrillation were scattered along the outcome (axis shown in Figure 2), whereas the patients who did not developed atrial fibrillation formed a cluster around this axis. The main reason for the variation along that axis came from the proteomic and metabolomic layers, suggesting that the shared biology among the susceptible patients was in the large part inflammatory and metabolic and not genetic or clinical: the results of integration did not betel about which layers to weight.

Table 3: Leading multi-omics latent factors and their association with post-operative atrial fibrillation

Factor	Dominant omics layer	Variance from that layer, %	Correlation with POAF	p
Factor 1	Proteomic	56	0.48	<0.001
Factor 3	Clinical	82	0.26	<0.001
Factor 2	Genomic	86	0.13	<0.001

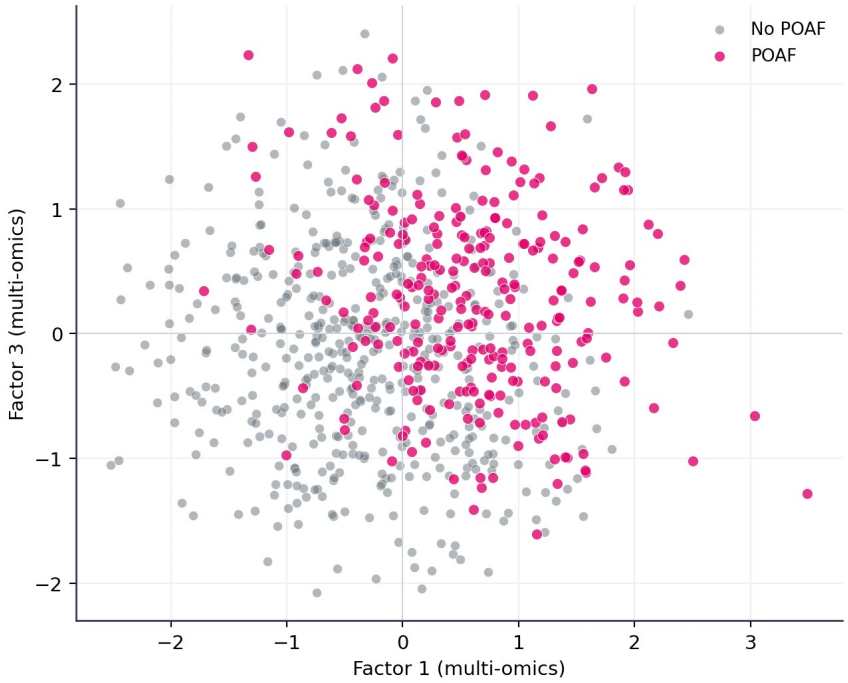


Figure 2: Patients positioned on the two leading multi-omics latent factors, coloured by outcome

3.4. The cross-omics POAF signature

The features most strongly associated with the outcome spanned the omics layers, as shown in Figure 3. They were led by inflammatory and cardiac-stress proteins, the natriuretic peptide NT-proBNP, growth differentiation factor 15 and interleukin-6, and were joined by metabolic markers, the ketone body acetoacetate, the microbial

co-metabolite trimethylamine-N-oxide and an acylcarnitine, together with glutamate (Wollert *et al.*, 2017; Tang *et al.*, 2017). That the signature is drawn from more than one layer is the point: it is the combination of an inflammatory proteomic signal and a metabolic one, neither sufficient alone, that marks the susceptible patient, which is why integrating the layers outperformed each of them.

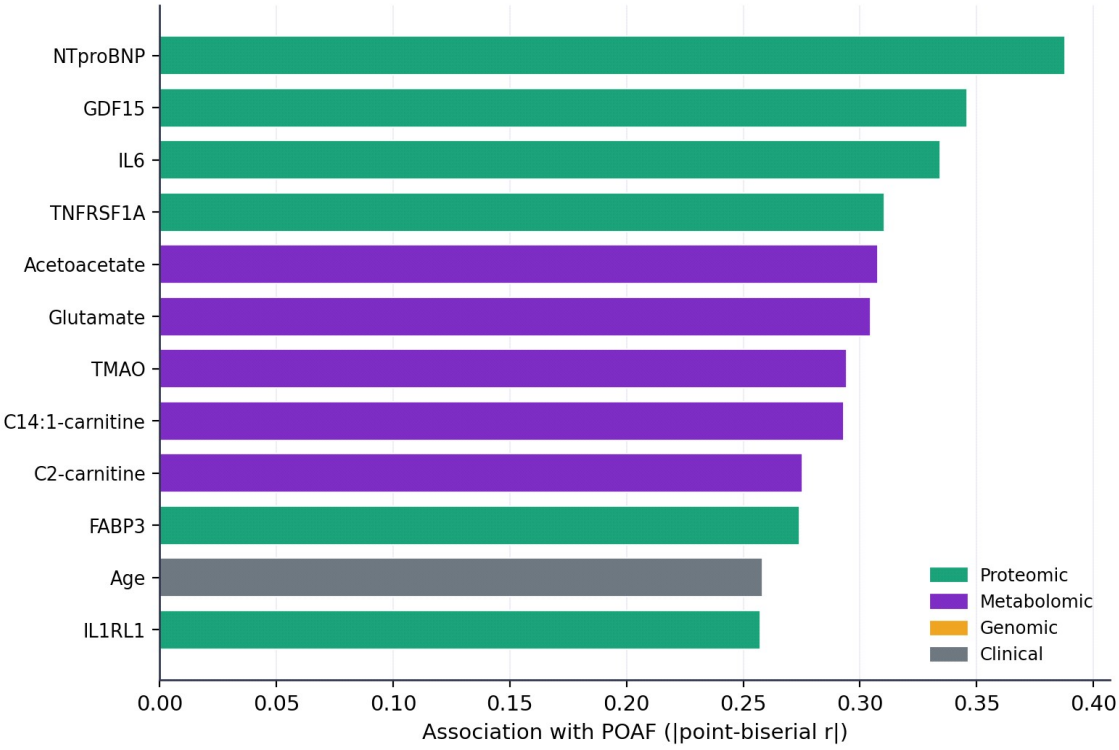


Figure 3: Leading features of the multi-omics signature of post-operative atrial fibrillation, coloured by omics layer

3.5. Calibration and incremental value

The integrated model was well calibrated across the range of predicted risk, and its advantage was evident not only in the area under the receiver operating characteristic curve but in the area under the precision-recall curve (0.71), the more searching measure given that only about a third of patients fibrillated. The incremental value over the clinical score was substantial, the integrated model raising the area from 0.66 to 0.82, a gain that in this setting corresponds to a materially better separation of the patients who will fibrillate from those who will not.

4. Discussion

Combining clinical, genomic, proteomic and metabolomic information led to a better pre-operative prediction of post-operative AF than any single layer alone, and better than the clinical score alone and there was a common post-op AF susceptibility inflammatory–metabolic axis. Risky signatures were taken from multiple levels simultaneously.

The main lesson which is the expected dividend of integration is the whole is greater than its parts. The proteome is the inflammatory state, the metabolome is the metabolic derangement, the genome is the inherited predisposition and the clinic is the structural substrate of each layer of the susceptibility. These facets are not totally independent, and, indeed, the stacked integration extracted more of the risk than each facet alone, and what the factor analysis revealed, is an axis chiefly driven by the proteomic and metabolomic ones. The low performance of the genomic score as compared to its contribution to the whole, demonstrates the same thing: a weak layer can still add when what it adds is complementary.

The biology that the signature refers to is consistent with the understanding of the disorder. The inflammation proteins (interleukin-6 and growth differentiation factor 15), the wall-stress marker NT-proBNP, and the

metabolic markers of ketosis and of microbial metabolism are all the molecular correlates of exactly that convergence of inflammation and metabolic disturbance, thought to be the trigger for post-operative atrial fibrillation. The occurrence of trimethylamine-N-oxide, a gut-microbial co-metabolite known to be involved in atrial remodelling and of the ketone body and acylcarnitine is indicative of metabolism as a yet underappreciated contributor, a metabolomic layer is needed to see.

The in the forefront of this value is the multi-omics approach in contrast to previous efforts to predict this arrhythmia. The discrimination of clinical scores composed only of age, atrial size and operatic type is modest, and incorporation of a sole biomarker or a small panel of such markers has improved such clinical scores modestly at best. On the present evidence, the reason is that no single marker—or even a narrow panel of single markers—can measure the several biologies that converge to produce the arrhythmia; it is the breadth of the multi-omics assessment from inflammation to metabolism to genetic predisposition and structure all at once and the integrated combination of them that provides the larger gain and the more useful. This is in line with the general rule that this kind of multi-omics integration works best in a biographically heterogeneous context: post-operative atrial fibrillation is such a disease.

For practice, a multi-omics assessment before the operation may better identify those most prone to fibrillate—in whom prophylaxis or closer monitoring of peri-operative risk factors, or a modification of the peri-operative strategy, should be focused—and in whom prevention would reap greatest benefit, more accurately than the clinical score. This is an ideal time to perform pre-operative patient profiling, when a patient's intervention date is known and the pre-operative period is a planned period of time. This does not address whether the use of the built-in prediction in turn improves outcomes, which would be the subject of a trial to determine direct effects of using it as part of care.

It is worth commenting on the selection for the integration of the layers used as this determined the outcome of the result. Two general approaches are: (1) early integration where the features of all layers are concatenated and modelled together, and (2) late integration where each layer is modelled independently and models integrated, e.g., by stacking, as done here. The benefit of late integration is that, in the presence of layers of very different kind and native size, each layer can model itself, and a flexible learner should model a large, high-dimensional native space, whereas a simple learner models a few low dimensional polygenetic scores; hence, their signals can be combined. These two complementary views—supervised and unsupervised—of structures obtained from the stacked model and factor analysis are that the stacked model integrated the layers into a predictor stronger than any single one of them, and that the factor analysis recovered the same inflammatory/metabolic axis constructed without supervision.

A subsequent finding of the factor analysis deserves special attention, as it addresses the for what of integration? The variation in the outcome-associated latent axis was mainly from the proteomic and metabolomic layers, but also included variation in the weak-genomic and clinical layers in the integrated predictor. That's the true essence of complementarity: you don't need to be an individually strong element to be a valuable one as long as it provides information not already found in the other elements. The reason that this case benefited from the addition of the polygen score, which captured an inherited predisposition that was not evident in the proteome or metabolome, and the clinical layer, which captured the structural substrate, was because each discriminated poorly on its own, but provided orthogonal information. The claim on this view is that the value of integration is not measured by the inclusion and stacking together of strong predictors, but by the integration of complementary predictors.

The four layers discussed here are not comprehensive of what may be integrated. Transcriptomic, epigenomic and microbiome data have been associated with atrial fibrillation and to the peri-operative inflammatory response and profiling patient cohorts across an even broader range of molecular data is underway. Here a further layer could easily be added: each would be modeled in its own way and placed on top of the current model, and the analysis of the data would allow for it in the search for the common axes. How useful additional layers might be for further prediction, or whether the additional layers would simply duplicate information already provided by the inflammatory and metabolic signal is an empirical question, but it has implications for the extent of a pre-operative, multi-omic assessment that would be necessary to support it. In light of the present results, the proteomic and metabolomic levels may be the most informative, and potentially a practical assessment could be made focusing on these levels.

6. Conclusion

A combined pre-op multi-omics assessment integrating clinical, genomic, proteomic and metabolomic data was able to predict post-op AF after cardiac surgery more accurately than any single data layer alone or the clinical score and revealed a common signature of inflammation and metabolism, across both proteins and metabolites, that distinguished patients at risk. The complementary information in several molecular layers could be integrated to recover risk that was not captured by any molecular layer alone; this has demonstrated the utility of multi-omics integration as a means to enhancing the ability to better stratify risk in the perioperative setting and suggests that the extent of the inflammatory and metabolic biology demonstrated through integration provides an opportunity for both predicting risk and, perhaps, preventing it.

Ethical approval

The study was approved by the Institutional Ethics Committee and was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from every participant.

Conflict of interest

The authors declare that no conflict of interest exists in relation to this work.

Funding

This work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability

The data supporting the findings of this study are available from the corresponding author on reasonable request.

References

Argelaguet, R., Velten, B., Arnol, D., Dietrich, S., Zenz, T., Marioni, J.C. *et al.* (2018). Multi-omics factor analysis: A framework for unsupervised integration of multi-omics data sets. *Mol Syst Biol*, 14(6): e8124.

Assarsson, E., Lundberg, M., Holmquist, G., Bjorkesten, J., Thorsen, S.B., Ekman, D. *et al.* (2014). Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS One*, 9(4): e95192.

Chen, T. and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-94.

Collins, G.S., Reitsma, J.B., Altman, D.G. and Moons, K.G. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Ann Intern Med*, 162(1): 55-63.

DeLong, E.R., DeLong, D.M. and Clarke-Pearson, D.L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 44(3): 837-45.

Dobrev, D., Aguilar, M., Heijman, J., Guichard, J.B. and Nattel, S. (2019). Postoperative atrial fibrillation: Mechanisms, manifestations and management. *Nat Rev Cardiol*, 16(7): 417-36.

Greenberg, J.W., Lancaster, T.S., Schuessler, R.B. and Melby, S.J. (2017). Postoperative atrial fibrillation following cardiac surgery: A persistent complication. *Eur J Cardiothorac Surg*, 52(4): 665-72.

Hindricks, G., Potpara, T., Dagres, N., Arbelo, E., Bax, J.J., Blomstrom-Lundqvist, C. *et al.* (2021). 2020 ESC guidelines for the diagnosis and management of atrial fibrillation. *Eur Heart J*, 42(5): 373-498.

Mariscalco, G., Biancari, F., Zanobini, M., Cottini, M., Piffaretti, G., Saccocci, M. *et al.* (2014). Bedside tool for predicting the risk of postoperative atrial fibrillation after cardiac surgery: The POAF score. *J Am Heart Assoc*, 3(2): e000752.

Picard, M., Scott-Boyer, M.P., Bodein, A., Perin, O. and Droit, A. (2021). Integration strategies of multi-omics data for machine learning analysis. *Comput Struct Biotechnol J*, 19: 3735-46.

Reel, P.S., Reel, S., Pearson, E., Trucco, E. and Jefferson, E. (2021). Using machine learning approaches for multi-omics data analysis: A review. *Biotechnol Adv*, 49: 107739.

Roselli, C., Chaffin, M.D., Weng, L.C., Aeschbacher, S., Ahlberg, G., Albert, C.M. *et al.* (2018). Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet*, 50(9): 1225-33.

Singh, A., Shannon, C.P., Gautier, B., Rohart, F., Vacher, M., Tebbutt, S.J. *et al.* (2019). DIABLO: An integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics*, 35(17): 3055-62.

Tang, W.H.W., Kitai, T. and Hazen, S.L. (2017). Gut microbiota in cardiovascular health and disease. *Circ Res*, 120(7): 1183-96.

Weng, L.C., Preis, S.R., Hulme, O.L., Larson, M.G., Choi, S.H., Wang, B. *et al.* (2018). Genetic predisposition, clinical risk factor burden, and lifetime risk of atrial fibrillation. *Circulation*, 137(10): 1027-38.

Wishart, D.S. (2019). Metabolomics for investigating physiological and pathophysiological processes. *Physiol Rev*, 99(4): 1819-75.

Wollert, K.C., Kempf, T. and Wallentin, L. (2017). Growth differentiation factor 15 as a biomarker in cardiovascular disease. *Clin Chem*, 63(1): 140-51.

Wolpert, D.H. (1992). Stacked generalization. *Neural Netw*, 5(2): 241-59.

Cite this article as: Suriya K. Wijaya (2026). Multi-omics integration for risk stratification of post-operative atrial fibrillation after cardiac surgery. *African Journal of Cardiology and Cardiovascular Medicine.* 2(1), 1-9. doi: 10.62587/AFJCCM.2.1.2026.1-9.