EBioMedicine 40 (2019) 23–24

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

EBioMedicine
Published by THE LANCET

Commentary

Artificially intelligent proteomics improves cardiovascular risk assessment

Siu Kwan Sze

School of Biological Sciences, Nanyang Technological University, Singapore

Cardiovascular disease (CVD) diagnosis, risk stratification, and treatment have improved significantly since the landmark Framingham Heart Study first defined key risk factors 50 years ago [1]. However, widespread use of indices such as the Framingham Risk Score (FRS) to guide patient management has not altered CVD status as the leading cause of mortality worldwide (still contributing to 1 in every 3 deaths in developed countries). This high burden of CVD persists due to the substantial amount of residual disease despite the use of anti-lipid, anti-hypertensive and anti-diabetic drugs for primary and secondary preventions. In addition, many patients who eventually suffer heart attack or stroke lack obvious premonitory symptoms and exhibit normal clinical risk factors [2,3]. Consequently, there is an urgent clinical need to better identify vulnerable patients who will go on to suffer major adverse cardiovascular events (MACE). Current blood plasma biomarkers of CVD risk such as C-reactive protein (CRP) and myeloperoxidase have so far failed to improve upon the discriminant value of FRS [3], and the latest large-scale study by the US Preventive Services Task Force found insufficient evidence that MACE can be better predicted by a range of novel physiological measures (including ankle-brachial index, CRP and coronary artery calcium score) [4]. Better strategies and more effective biomarkers of clinical course will therefore be required in order to develop better prognostic tests and interventions for vulnerable patients at risk of MACE.

CVD is a complex disorder that results from the interplay of multiple environmental and genetic factors, hence the molecular basis of this disease is still not fully understood. In recent decades, discovery-led proteomics and hypothesis-based studies have identified a variety of biomarkers that correlate with various pathological features of CVD, and may therefore prove useful for assessing risk in specific subtypes of disease. However, single biomarkers of individual clinical characteristics are unlikely to significantly improve on FRS-based risk assessment across the wider CVD population. More likely is that a 'biosignature' comprising several biomarkers of multiple pathological features will allow improved risk stratification across a diverse CVD cohort, although previous work in the area has yet to yield consensus on the clinical value of this approach. This could in-part be due to previous studies being guided by specific hypotheses that focused on a limited repertoire of potential biomarkers, and were therefore unable to assess the full complexity of CVD pathogenesis. In contrast, discarding *a priori* hypotheses in favor of big data discovery-driven approaches with machine learning/artificial intelligence techniques could potentially identify more effective biosignatures capable of major advances in clinical risk assessment [5].

In this issue of *EBioMedicine*, Bom et al. report two plasma protein signatures for risk stratification in a cohort of 203 patients with suspected coronary artery disease [6]. The authors used targeted proteomics to quantify 332 known CVD and inflammation-associated proteins in patient plasma samples and used a machine learning method to identify a subset of 35 proteins that predicted high-risk atherosclerotic plaque (defined by coronary artery calcium score and computed tomography angiography), as well as a separate biosignature of 34 plasma proteins that predicted lack of atherosclerotic pathology. While these findings are currently restricted to a single small cohort and remain unvalidated, prospective studies will now be able to assess how these biosignatures correlate with MACE risk in larger patient groups and determine prognostic power in diverse populations. This study demonstrates that proteomics coupled with machine learning is now an essential method for uncovering signatory profiles of distinct clinical outcomes in complex disorders such as CVD. This approach will allow biomarker research to progress beyond correlating soluble protein levels with disease state to also consider the roles of extracellular vesicles (EVs), protein posttranslational modifications (PTMs) and degenerative protein modifications (DPMs) in tandem. Indeed, several recent studies have shown that immune cells, platelets and endothelial cells secrete EVs that contain diverse proteins, nucleic acids, and metabolites which can critically influence the clinical course of CVD [7]. Using artificial intelligence approaches to probe this complexity will likely uncover important new features of CVD pathophysiology that could not be predicted by conventional techniques alone. For example, DPMs produced by spontaneous chemical reactions can cause age-dependent damage to critical host proteins that lead to progressive loss of essential functions [8,9], but advanced proteomics studies recently revealed that 'protein aging' can also confer paradoxical 'gain-of-function' changes by generating isoDGR motifs in the vascular bed matrix [10]. These *de novo* integrin-binding structures promote leukocyte recruitment and stimulate expression of pro-inflammatory cytokines in the blood vessel wall, and may therefore represent important new biomarkers of CVD risk that could not be predicted based on earlier research. Future studies will therefore be able to use advanced

DOI of original article: https://doi.org/10.1016/j.ebiom.2018.12.033.

E-mail address: sksze@ntu.edu.sg.

https://doi.org/10.1016/j.ebiom.2019.01.014

proteomics methods to combine protein expression levels and PTM data with DPM profiles and analyses of EV cargo, thereby defining more effective biosignatures of clinical course in multifactorial disorders such as CVD.

Acknowledgements

This work is in part supported by the Singapore National Medical Research Council (NMRC/OFIRG/0003/2016), and Singapore Ministry of Education (MOE2016-T2-2-018 and MOE2018-T1-001-078).

Disclosure

The author disclosed no conflicts of interest.

References

[1] Dawber TR, Kannel WB, Revotskie N, Stokes 3rd J, Kagan A, Gordon T. Some factors associated with the development of coronary heart disease: six years' follow-up experience in the Framingham study. Am J Public Health Nations Health 1959;49:1349–56.
[2] Fernandez-Friera L, Fuster V, Lopez-Melgar B, et al. Normal LDL-cholesterol levels are associated with subclinical atherosclerosis in the absence of risk factors. J Am Coll Cardiol 2017;70(24):2979–91.
[3] Shah PK. Screening asymptomatic subjects for subclinical atherosclerosis: can we, does it matter, and should we? J Am Coll Cardiol 2010;56(2):98–105.
[4] Curry SJ, Krist AH, Owens DK, et al. Risk assessment for cardiovascular disease with nontraditional risk factors: US preventive services task force recommendation statement. JAMA 2018;320(3):272–80.
[5] Motwani M, Dey D, Berman DS, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. Eur Heart J 2017;38(7):500–7.
[6] Bom MJ, Knaapen P, et al. Predictive value of targeted proteomics for coronary plaque morphology in patients with suspected coronary artery disease. EBioMedicine 2019. https://doi.org/10.1016/j.ebiom.2018.12.033.
[7] Sluijter JPG, Davidson SM, Boulanger CM, et al. Extracellular vesicles in diagnostics and therapy of the ischaemic heart: Position paper from the working group on cellular biology of the heart of the european society of cardiology. Cardiovasc Res 2018;114(1):19–34.
[8] Truscott RJW, Schey KL, Friedrich MG. Old proteins in man: A field in its infancy. Trends Biochem Sci 2016;41(8):654–64.
[9] Gallart-Palau X, Serra A, Sze SK. Uncovering neurodegenerative protein modifications via proteomic profiling. Int Rev Neurobiol 2015;121:87–116.
[10] Dutta B, Park JE, Kumar S, et al. Monocyte adhesion to atherosclerotic matrix proteins is enhanced by Asn-Gly-Arg deamidation. Sci Rep 2017;7(1):5765.