NOVEL BIOMARKERS FOR PREDICTING CLINICAL OUTCOMES IN CANCER PATIENTS

Identifying novel biomarkers using proteomics to predict cancer-associated thrombosis

Maria J. Fernandez Turizo,1 Rushad Patell,2 Jeffrey I. Zwicker3,4

1Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA; 2Division of Medical Oncology and Hematology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA; 3Department of Medicine, Hematology Service, Memorial Sloan Kettering Cancer Center, New York, NY; 4Weil Cornell Medical College, New York, NY, United States

ABSTRACT

Comprehensive protein analyses of plasma are made possible by high-throughput proteomic screens, which may help find new therapeutic targets and diagnostic biomarkers. Patients with cancer are frequently affected by venous thromboembolism (VTE). The limited predictive accuracy of current VTE risk assessment tools highlights the need for new, more targeted biomarkers. Although coagulation biomarkers for the diagnosis, prognosis, and treatment of VTE have been investigated, none of them have the necessary clinical validation or diagnostic accuracy. Proteomics holds the potential to uncover new biomarkers and thrombotic pathways that impact the risk of thrombosis. This review explores the fundamental methods used in proteomics and focuses on particular biomarkers found in VTE and cancer-associated thrombosis.

Introduction

Venous thromboembolism (VTE) represents a significant global public health issue, impacting roughly 10 million individuals annually worldwide and contributing to over 3 million annual fatalities. Among patients with cancer, thrombotic events are highly prevalent, with active cancer accounting for 20% of the overall incidence of VTE. The annual incidence of VTE in patients with cancer is 5-20% in higher-risk malignancies, compared to 0.1% in the general population. Clinical biomarkers for the diagnosis, risk prediction, recurrence estimation, and response to treatment in cancer-associated thrombosis (CAT) are limited.

D-dimer as a marker of endogenous fibrinolysis, has found utility as a valuable biomarker in clinical practice for the diagnosis of VTE. Nevertheless, it lacks specificity, such that is largely used as a diagnostic assay to exclude VTE, primarily due to its strong negative predictive value. Although various additional biomarkers, including P-selectin, tissue factor (TF), microRNAs, among others, have been investigated, none have been validated sufficiently for routine application in clinical practice.

Predicting VTE recurrences continues to pose challenges. D-Dimer in combination with clinical and genetic risk factors has been applied to help predict which patients will develop recurrent VTE following a course of therapeutic anticoagulation. Various studies have indicated that, following an initial spontaneous VTE, patients with low D-Dimer levels have a low risk of VTE recurrence upon discontinuation of anticoagulation. Conversely, patients experiencing a provoked VTE with elevated D-Dimer levels after discontinuing anticoagulation therapy, have an increased risk for VTE recurrence.

Biomarkers for thrombosis prediction in cancer

Standard cutoffs for D-Dimer have limited specificity, particularly in cancer patients in which D-Dimer levels are often increased at baseline. In cancer patients, higher levels of D-Dimer,
above the 75th percentile, have been found to correlate with an increased risk of VTE. A rising D-Dimer level over time has also been found to be predictive of VTE in the cancer population. The Khorana score is useful for VTE prediction in ambulatory cancer patients, with a high negative predictive value (≥80%). The score incorporates pre-treatment platelet count and leukocyte count, hemoglobin level, cancer type, and body mass index (BMI). The positive predictive value of a higher-risk Khorana score is approximately 10%.

Different models have been developed to improve the accuracy of thrombosis prediction with mixed results, using various cutoffs for D-Dimer, and the addition of biomarkers to the Khorana score. The Vienna Cancer and Thrombosis study score added D-Dimer and soluble p-selectin to the Khorana score factors, with an accurate prediction of VTE. The PROTECHT removed BMI from the prediction model and included chemotherapy. Additional prediction models, including ONKOTEV, COMPASS-CAT, Tic-ONCO, and IMPEDE, among others, have attempted to enhance the diagnostic accuracy of VTE in patients with malignancies. These models integrate various factors such as various types of malignancies, cancer stages, genetic risk factors, and D-Dimer levels.

Considering D-Dimer’s low positive predictive value, limited specificity, and modest discriminatory ability in cancer patients, there is a need for novel specific biomarkers to more effectively exclude VTE in this population. Additionally, traditional clinical models, including the Wells’ or Geneva scores, show limited efficacy in ruling out VTE in cancer patients. The mere presence of a comorbid malignancy elevates the clinical probability of VTE, requiring imaging for the majority of cancer patients to effectively exclude thrombosis.

Measuring the proteome

Based on the modest diagnostic and predictive accuracy of available coagulation biomarkers in VTE, the question remains whether measurement of other circulating plasma proteins offers clinical benefit. Proteomic screens are promising not only for discovering novel biomarkers for VTE in cancer but also for enhancing our understanding of the underlying pathophysiology of thrombus formation and the complex interplay among various prothrombotic factors, such as chemotherapy, immune response, and underlying malignancies. Various technologies are utilized to measure proteins in tissues, serum, or plasma, including highly optimized single protein assays, mass spectrometry (MS), and affinity-based assays. Below is a summary of proteomic methodologies and observations to date pertaining to VTE.

Mass spectrometry proteomics

MS is a technique used for the identification and quantification of proteins within a sample, allowing for customized measurements of specific targets of interest. It provides valuable insights into the structure, function, and composition of the proteome across diverse biological systems. The process involves ionizing peptides generated by proteolysis, a step accomplished through methods like electrospray ionization, surface-enhanced laser desorption ionization, and matrix-assisted laser desorption ionization. Subsequently, an electric or magnetic field is used to separate the ionized peptides in a mass analyzer, based on their mass-to-charge ratios (m/z). Tandem MS enhances the confidence of peptide identification by using molecules that have undergone prior fragmentation, performing further fragmentation, and isolation in a secondary mass analyzer.

Common combinations of mass analyzers and ionization methods include matrix-assisted laser desorption ionization, and liquid chromatography (LC) coupled with tandem MS. To aid peptide identification, various processes are typically employed, including sample enrichment, fractionation, depletion, and labeling. These methods promote protein separation, enhance detection sensitivity, and facilitate the identification of less abundant plasma proteins. Following peptide identification, specialized software tools and search algorithms are used to identify the parent protein in online databases. To enhance result confidence, statistical methods and secondary proteomics techniques are employed for validation. Furthermore, multiplexed MS allows for the simultaneous analysis of multiple samples in a single measurement, employing various methods such as isobaric labeling or label-free quantification.

Various methods deploy MS techniques for proteomic screening, including shotgun proteomics and targeted strategies. Shotgun proteomics, frequently employed in discovery studies, indirectly measures entire proteomes by analyzing peptides produced through the enzymatic breakdown of intact proteins. Analyzing complex samples such as plasma, poses a particular challenge in shotgun proteomics due to the differential abundance of proteins. In contrast, targeted proteomics approaches provide an alternative to the shotgun method, employing specific ions to identify a predefined set of peptides. Targeted approaches demonstrate high sensitivity, specificity, and reproducibility, proving effective for protein identification in diverse samples, and excluding the need for affinity reagents.

MS proteomics offers some advantages in comparison to conventional affinity-based assays (Table 1), such as its high specificity and ability to process large sample volumes. In addition, MS allows for the examination of post-translational modifications and characterization of isoforms, enhancing its versatility in protein analysis. On the contrary, limitations of the technique include a constrained ability to detect low abundance proteins in plasma, a limited dynamic range potentially hindering a comprehensive assessment of the proteome. MS encounters difficulties analyzing large, hydrophobic proteins and complex samples. It is also expensive and requires skilled operators and advanced instrumentation. Notably, the throughput capacity of conventional MS is comparatively lower than that of affinity-based assays. This disparity contributes to the weak correlation observed between MS and affinity-based platforms.

Mass spectrometry in venous thromboembolism proteomics

The literature on proteomics techniques for identifying clinical biomarkers in thrombosis is heterogeneous and limited, ranging from the analysis of complex plasma samples to specific protein analyses within distinct cellular subsets. Table 2 summarizes studies focusing on the analysis of plasma samples using proteomics in VTE. Zhang et al. used matrix-assisted laser desorption ionization MS analysis in patients with acute pulmonary embolism (PE) and healthy controls. After validation with ELISA, only haptoglobin was associated with PE. Various studies have associated haptoglobin...
globin levels with VTE. It appears that acute PE without pulmonary hypertension induces haptoglobin, but severe PE causes the haptoglobin concentration to decrease in proportion to the severity of the pulmonary hypertension.38,39

Han et al. profiled proteins from 13 plasma samples using MS and 32 plasma samples using antibody-based-assay proteomics. Samples were obtained from two independent case-control studies of patients with high-risk PE, non-high-risk PE, and healthy controls. Serum amyloid A-1 (SAA1), calprotectin (S100A8), tenascin-C (TNC), gelsolin (GSN), and histidine-rich glycoprotein (HRG), were differentially expressed in patients with PE and or in high-risk PE, in comparison to healthy controls.40

Tandem MS has also been used to analyze plasma samples from patients with VTE. Jensen et al. found that Transthyretin, vitamin K-dependent protein Z, and protein/nucleic acid deglycosylation with VTE and healthy controls.41

Affinity proteomics

Affinity proteomics employs binding agents to serve as probes for the targeted detection of proteins. Binding agents include antibodies or aptamers, single-stranded DNA or RNA molecules designed to selectively bind specific targets.42 Diverse platforms use affinity proteomics for the large-scale study of biomarkers, including antibody-based suspension bead arrays, proximity extension assays, surface arrays, and the aptamer-based Soma Scan assay.27,43,44 All platforms enable multiplexed profiling of proteins, enabling simultaneous analysis of samples within a single experiment.45

Affinity-based assays offer several notable advantages, including a wide dynamic range, multiplexing capability, high specificity that minimizes cross-reactivity with non-target proteins, and versatility across various sample types. Limitations of affinity-based proteomics include alterations between targets and binding agents. For instance, structural and conformational changes in proteins, nonspecific protein binding, missense mutations, single nucleotide polymorphisms, and differential splicing can disturb the interaction between binding epitopes and targets.44,46 Mutations within a gene's coding regions induce alterations in the amino acid sequence of the associated protein. Changes that occur within the epitope target region may significantly impact the protein’s binding affinity.46 To address this issue, various methods have been developed to validate the target of the binding agent.47

Table 1. Comparison between mass-spectrometry and affinity-based assays.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mass-spectrometry</th>
<th>Proteomics technologies</th>
<th>Affinity-based assays</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein detection</td>
<td>Measures m/Q of peptide fragments</td>
<td>Uses antibodies or aptamers that bind to proteins</td>
<td></td>
</tr>
<tr>
<td>Protein quantification</td>
<td>Provides absolute quantification of proteins, with an average of 10 peptides per protein</td>
<td>Provides relative quantification but not absolute quantification of proteins</td>
<td></td>
</tr>
<tr>
<td>Dynamic range</td>
<td>Lower sensitivity for the detection of low abundance proteins (&lt;10 ng/mL)</td>
<td>Detects high and medium abundance proteins.</td>
<td>Higher sensitivity for the detection of low abundance proteins</td>
</tr>
<tr>
<td>Specificity</td>
<td>High degree of specificity, Not restricted to a predefined set of targets</td>
<td>Suitable for use in diverse species and across a wide range of sample types</td>
<td>Specifictiy varies according to the assay type and degree of cross-reactivity</td>
</tr>
<tr>
<td>Characterization of structural modifications</td>
<td>Detects conformational changes in proteins, diverse protein isoforms, and post-translational modifications</td>
<td>Unable to detect protein isoforms, post-translational modifications, and other proteoforms</td>
<td>Alteration of binding sites leads to a decrease in specificity</td>
</tr>
<tr>
<td>Throughput capacity</td>
<td>Traditional MS methods have limited throughput capacity (low to moderate).</td>
<td>Affinity enriched or affinity selection MS methods are used to overcome this challenge</td>
<td>High throughput capacity using various affinity reagents</td>
</tr>
<tr>
<td>Reproducibility</td>
<td>Lower reproducibility,Modest intra-assay variation</td>
<td>Higher reproducibility,Low intra-assay variation</td>
<td></td>
</tr>
<tr>
<td>Sample volume</td>
<td>Large samples (30-100 μL)</td>
<td>Immunoaffinity arrays (PEA) 1-100 μL</td>
<td>Immunoaffinity arrays (PEA) 100 proteins</td>
</tr>
<tr>
<td>Multiplexing</td>
<td>All proteins in the sample (10-&gt;5000)</td>
<td>Aptamer-based arrays: 65 μL</td>
<td>Aptamer-based arrays: &gt;1300 proteins</td>
</tr>
</tbody>
</table>

m/Q, mass-to-charge ratio; MS, mass spectrometry; PEA, proximity extension assays.
### Table 2. Proteomic studies in venous thromboembolism.

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Sample</th>
<th>Cohort</th>
<th>Comparative groups</th>
<th>Methods</th>
<th>Results</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brzelius et al. (^{33}) 2016</td>
<td>Plasma</td>
<td>Patients from the VE BIOS and FARIVE studies</td>
<td>VEBIOS VTE (n=88) Healthy controls (n=85) Replication study FARIVE VTE (n=580)</td>
<td>IC-MS ELISA Bead arrays 755 antibodies targeting 408 proteins</td>
<td>VWF and PDGFB levels were significantly higher in patients with VTE *Results were verified using patients from the FARIVE study</td>
<td>VWF (P=0.001) PDGFB (P=0.002) Pearson’s correlation between studies: VE BIOS:0.42 FARIVE:0.26</td>
</tr>
<tr>
<td>Jensen et al. (^{34}) 2018</td>
<td>Plasma</td>
<td>Patients from the Tromsø Study</td>
<td>VTE (n=100) Healthy controls (n=100)</td>
<td>TMT LC-MS</td>
<td>Strongest biomarkers for the development of VTE: -Transhthrelin -Protein Z (ProZ) -Protein/nucleic acid deglycase (DJ-1)</td>
<td>Transthyrelin P=0.00015 ProZ P=0.0018 DJ-1 P=0.0055</td>
</tr>
<tr>
<td>Razzaq et al. (^{59}) 2021</td>
<td>Plasma</td>
<td>1388 Patients with DVT with or without PE from the MARTHA and EOVT studies</td>
<td>MARTHA PE (n=95) DVT (n=1105) DVT+PE (n=188) Verification study EOVT PE (n=143) DVT (n=196)</td>
<td>1. SBA in combination with Machine learning methods-ANN model 2. Application of the LIME algorithm 3. GWAS conducted on the LIME estimate</td>
<td>PLXNA4 was identified as a susceptibility locus for isolated PE phenotype. Homozygote carriers for the rs1424597-A allele were more frequently observed in PE than in DVT patients *Results were verified using patients from the EOVT study</td>
<td>GWAS on the LIME estimate (rs1424597): P=5.3×10^-7 at the PLXNA4 locus Homozygote carriers -isolated PE phenotype vs DVT MARTHA (2% vs. 0.4%) P=0.005 EOVT (3% vs. 0%) P=0.013</td>
</tr>
<tr>
<td>Ten Cate et al. (^{49}) 2021</td>
<td>Plasma</td>
<td>532 Patients from the GMP-VTE study</td>
<td>GMP-VTE PE (n=96) DVT (n=160) DVT+PE (n=276) Verification study Gutenberg Health study (n=5778)</td>
<td>96-plex Olink panels (cardiometabolic, cardiovascular II and III, inflammation, and immune response) Proteomics in combination with machine learning: LASSO-regularized regression models</td>
<td>Prognostic proteins for the development of primary isolated PE in comparison to DVT or DVT+PE: -Interferon-γ -GDNF -Interleukin-15Rα *Results were verified using patients from the Gutenberg health study</td>
<td>HR per SD increase -Interferon-γ HR (1.34 95% CI, 1.23-1.45; P=0.0001) -GDNF HR (0.40 5% CI, 0.29-0.55; P&lt;0.0001) -Interleukin 15Rα HR (0.55 95% CI, 0.43-0.71; P&lt;0.0001)</td>
</tr>
<tr>
<td>Han et al. (^{43}) 2021</td>
<td>Plasma</td>
<td>Patients with PE and healthy controls from two case control studies</td>
<td>Discovery MS analysis: high-risk PE (n=3) non-high-risk PE (n=6) healthy controls (n=4) Antibody array analysis high-risk PE (n=10) non-high-risk PE (n=10) healthy controls (n=12)</td>
<td>MS Antibody array proteomic technology ELISA</td>
<td>Differentially expressed proteins in patients with PE/High-risk PE: -SAA1 -S100A8 -Tenascin-C(TNC) -Gelsolin -HRG *Results were verified using an independent cohort of 76 patients</td>
<td>AUC for PE diagnosis: -SAA1 Cut-off:1.26 μg/ml (AUC 0.882) -S100A8 Cut-off:1.19 μg/ml (AUC 0.788) -TNC Cut-off:12.62 μg/ml (AUC 0.720) AUC for High-risk PE diagnosis: -S100A8 Cut-off:1.7 μg/ml (AUC 0.773) -TNC Cut-off:17 μg/ml (AUC 0.720)</td>
</tr>
</tbody>
</table>

*Results were verified using patients from the FARIVE study

To be continued on next page
Proximity extension assays in venous thromboembolism proteomics

The literature on PEA studies in VTE proteomics represents a limited yet evolving landscape, marked by significant variability in methodologies across studies. While some investigations identify specific protein associations with VTE phenotypes, the overall heterogeneity in approaches emphasizes the need for further standardization and larger-scale studies. Below are a few highlighted key studies.

Ten Cate et al. identified 5 proteins specifically associated with an isolated PE phenotype, compared with deep vein thrombosis.

Table 2. Continued from previous page.

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Sample</th>
<th>Cohort</th>
<th>Comparative groups</th>
<th>Methods</th>
<th>Results</th>
<th>Parameters (cutoff, sensitivity, specificity, significance, correlation, HR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhang et al., 2014</td>
<td>Plasma</td>
<td>Patients with PE and matched healthy controls</td>
<td>18 patients PE (n=9)</td>
<td>2DE MALDI-TOF MS ELISA</td>
<td>Haptoglobin was overexpressed in the serum of PE patients.</td>
<td>Haptoglobin cut-off: 256.74 mg/l (95% CI, 0.622–0.906) P&lt;0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Healthy controls: (n=9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Verification study 48 patients PE (n=24)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Healthy controls (n=24)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memon et al., 2018</td>
<td>Plasma</td>
<td>357 patients with suspected DVT from a prospective multicenter (7 centers) management study in southern Sweden</td>
<td>90 patients included Confirmed acute DVT (n=45)</td>
<td>Olink Panel Cardiovascular III</td>
<td>Proteins significantly associated with VTE:</td>
<td>-P-Selectin AUC 0.84 (95% CI 0.76-0.92) P=0.000001 -TF pathway inhibitor TFPI AUC 0.74 (95% CI 0.64-0.85) P=0.00001 -VWF AUC 0.77 (95% CI 0.67-0.87) P=0.00001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Healthy matched controls (n=45)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Verification studies:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-PEA -DVT -MALDI-TOF MS ELISA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-Proteins significantly associated with VTE:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-P-Selectin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-TF pathway inhibitor TFPI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-VWF</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-Osteopontin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-Bleomycin hydrolase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-ST2 protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iglesias et al., 2023</td>
<td>Plasma</td>
<td>Patients from the VEBIOS study</td>
<td>VTE (n=144)</td>
<td>SBA</td>
<td>Complement factor H Diagnosis of acute VTE associated with 1 SD increase of CFHR5 concentration:</td>
<td>OR 2.54 (95% CI 1.52-4.66) P=1.05E-03</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Healthy controls (n=140)</td>
<td>LC-MS/MS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Verification studies:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-DFW-VTE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-FARIVE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-RETROVE -MARTHA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2DE, two-dimensional gel electrophoresis; ANN, artificial neural networks; DVT, deep venous thrombosis; DFW-VTE, Swedish Karolinska age-adjusted D-dimer study; FARIVE, French multicenter case-control study; GDNF, glial cell-line derived neurotrophic factor; GMP-VTE, genotyping and molecular phenotyping in venous thromboembolism study; GWAS, genome-wide association study; HR, hazard ratio; HRG, histidine-rich glycoprotein; IC, immunocapture; LC, liquid chromatography; LIME, local interpretable model-agnostic explanations; MARTHA, Marseille Thrombosis Association study; MS, mass spectroscopy; RETROVE, Riesgo de Enfermedad Tromboembólica Venosa study; SAA1, serum amyloid A-1; SBA, suspension bead array; SD, standard deviation; PE, pulmonary embolism; TF, tissue factor; TMT, tandem mass tag; VEBIOS, venous thromboembolism biomarker study; VTE, venous thromboembolism; VWF, Von Willebrand Factor; DJ-1, deglycase; PDGFβ; platelet-derived growth factor subunit B.
bosis (DVT) or DVT-associated PE phenotypes. Using 5 PEA panels, 3 proteins (interferon-γ, glial cell-line derived neurotrophic factor, and interleukin-15Rα) were found to be differentially expressed in VTE patients. Ligation of the inferior vena cava to induce DVT in mice, demonstrated that intrathrombotic levels of interferon-γ were progressively elevated as the post ligation interval extended. In addition, Interleukin15 complexes have a well-established role in cardiovascular disease, participating in inflammatory pathways and coronary thrombosis.

In the context of DVT, Memon et al. employed a single PEA panel to profile proteins in patients with acute DVT and matched controls. The study identified 7 proteins significantly associated with VTE, including p-Selectin, TF pathway inhibitor, VWF, and platelet rolling and adhesion, enhances TF expression in monocytes, and instigates the release of procoagulant substances and platelet activation. The role of transferrin receptor protein 1, osteopontin, monocytes, and ST2 in thrombosis remains under investigation.

PLXNA4 was identified as a new susceptibility locus for PE. The MARTHA study involved targeted screening, involving the immobilization of antibodies into microscopic beads. Each bead is conjugated with an antibody that interacts with proteins from a biological sample and creates complexes that can be quantified. Unique fluorescent labels are often attached to allow the identification of the complexes. Bead-based assays have a high throughput capacity and high multiplexing ability. They have been applied for various purposes, including the detection of cytokines, auto-antibodies, the analysis of monoclonal antibodies, and biological warfare agents.

Bead-based assays

Bead-based assays are antibody-based methods for proteomic screening, involving the immobilization of antibodies into microscopic beads. Each bead is conjugated with an antibody that interacts with proteins from a biological sample and creates complexes that can be quantified. Unique fluorescent labels are often attached to allow the identification of the complexes. Bead-based assays have a high throughput capacity and high multiplexing ability. They have been applied for various purposes, including the detection of cytokines, auto-antibodies, the analysis of monoclonal antibodies, and biological warfare agents.

Bead-based assays in venous thromboembolism proteomics

The VEREMA affinity proteomics study assessed plasma samples using bead arrays obtained from patients with VTE and matched healthy controls. A set of 408 proteins, selected for their known involvement in the coagulation cascade, expression in endothelial cells, and associations with cardiovascular disease and inflammation pathways, served as targets. The findings were then compared to plasma samples from the French FARIVE study for replication, ultimately confirming the independent associations of VWF and platelet-derived growth factor subunit B (PDGFB) with VTE. PDGF is expressed in endothelial cells and platelets, and elevated levels are associated with an increased risk of thrombosis.

Various studies have used bead-based assays to identify biomarkers that are able to distinguish between PE and DVT. Razzaq et al. analyzed plasma samples of patients with VTE from the Marseille Thrombosis Association study (MARTHA) study using a machine learning framework employing an artificial neural network approach to integrate plasma proteomics with genetic data. The MARTHA study involved targeted affinity proteomics using suspension bead assay technologies. PLXNA4 was identified as a new susceptibility locus for PE. PLXNA4 plays an important role in pathways related to thrombosis, stimulating TNF-α and IL-6 production in macrophages. Its ligand SEMA3, is known to promote vascular remodeling and regulate platelet aggregation. It has been strongly associated with various lung function markers but its precise association with PE is still under study.

Complement factor H-related 5 (CFHR5) protein represents a potential diagnostic and or risk predictive biomarker for VTE. Suspension bead arrays were used to analyze plasma samples obtained from patients in the VEBIOS study. Elevated levels of CFHR5 were associated with increased thrombin generation and platelet activation in vitro. Notably, the association between CFHR5 and VTE was also reported in a cohort of patients with COVID-19 infection.

Aptamer-based assays

Aptamers are nucleotide-based agents with protein affinity. Large nucleotide sequences are mixed with target peptides or proteins for binding. A commercial platform based on a large library of synthetic oligonucleotide ligands was developed by Somascan. The specificity of the technique can be limited by cross-reactivity among agents. Somascan aptamers are modified with aromatic benzyl side chains to reduce cross-reactivity. Aptamer-based assays have a high sample throughput and multiplexing capacity, with a wide dynamic range, and no toxic or immunogenic potential.

There are limited studies to date evaluating proteins through the Somascan platform and the development of thrombosis. In a study of 59 critically ill adolescents using data obtained from the Somascan platform, 9 patients developed incident DVT. Higher levels of thrombin-antithrombin complexes and lower levels of factor XIII were associated with DVT. In addition, CD36, macrophage inhibitory cytokine-1, and erythropoietin receptor were marginally associated with DVT.

Comparative analysis of proteomics techniques

Comparing different proteomic platforms such as MS, antibody, and aptamer, affinity-based assays have generally demonstrated limited correlation. Although consistent and comparable outcomes across different platforms are lacking.

The analysis of 173 human blood plasma samples using both MS-based platforms and PEA (Olink), identified 35 proteins common to both techniques. The two MS platforms demonstrated a strong correlation coefficient exceeding 0.5 for 23 of these 35 proteins. However, across all three platforms, including PEA and MS, only 6 out of the 35 proteins exhibited a correlation coefficient exceeding 0.5.

Various investigations have found a weak correlation between PEA (Olink) and the SomaScan platforms. However, studies have been constrained by a limited number of analyzed proteins and a small sample size. For instance, in a comparative study of 27 healthy individuals and 27 with acute VTE, there was a poor agreement for 8 common coagulation proteins including D-dimer and fibrinogen. In addition, a large-scale plasma proteomics study comparing the United Kingdom Biobank Olink (PEA) and Iceland SomaScan platforms, revealed a modest Spearman correlation between both techniques.
Proteomics in cancer-associated thrombosis

CAT exhibits distinctive features that set it apart from other types of VTE, including differences in risk factors, pathophysiology, and management strategies. Central to its pathogenesis is the pivotal role of TF, a key player in cancer progression and CAT.\(^7\) TF induces the activation of platelets and the coagulation cascade. Its release into circulation occurs within TF-positive extracellular vesicles. Notably, certain tumor types, including pancreatic, ovarian, brain, and cervical cancers, manifest elevated levels of TF, with potential correlations to specific oncogenic gene mutations, angiogenesis, and tumor histological grade.\(^7\)\(^3\)\(^4\) Procoagulant proteins such as plasminogen activator inhibitor 1, podoplanin, and protein disulfide isomerase have also been implicated in CAT. Table 3 provides a summary of proteomic plasma biomarkers evaluated in CAT.

Differential proteomic expression in various malignancies

Proteomic investigations include differential protein expression across various malignancies as they relate to CAT. In a study of patients with lung (N=30, 15 with VTE) and pancreatic cancer (N=30, 15 with VTE) using LC-MS, there were distinct differential expression patterns of immunoglobulin-derived proteins and tetranectin in cancer patients with and without VTE. Particularly noteworthy was the absence of overlap between lung and pancreatic cancer, emphasizing the nuanced variations in mechanisms and proteins based on the primary malignancy site.\(^7\)\(^8\) Cancer-derived immunoglobulins are highly expressed in cancer cells and mediate multiple processes in cancer progression, coagulation, and inflammation, including activation of platelet aggregation.\(^7\)\(^6\) Furthermore, the analysis of plasma samples from 20 patients with non-small cell lung cancer (NSCLC) and VTE, and 15 NSCLC patients without VTE, demonstrated differential expression of 5 proteins (SAA1, S100A8, lipopolysaccharide binding protein, haptoglobin, and lactate dehydrogenase B) in VTE patients.\(^7\)

The platelet proteome

Platelets play a crucial role in cancer biology and CAT. Research has indicated that the platelet proteome exhibits variations based on the primary site of malignancy. For example, in a MS proteomics study involving patients with brain cancer, lung cancer, and healthy controls, while the platelet proteome remained unaltered in brain cancer, distinctive modifications and differential expression of proteins were observed in patients with lung cancer when compared to the healthy control group.\(^7\)\(^9\) Furthermore, a separate study involving the platelet proteome of 9 individuals with diverse malignancies found that the platelet proteome was affected not only by the type of primary malignancy but also by the oncological treatment.\(^7\)\(^9\)

Extracellular vesicles

Extracellular vesicles (EVs) facilitate the interaction between cancer cells, platelets, and the vascular system. In the context of CAT, cancer cells release EVs containing diverse bioactive substances, such as proteins, nucleic acids, and lipids. These EVs contribute to the hypercoagulability observed in cancer patients. Specifically, EVs released by cancer cells can activate platelets, inducing platelet aggregation and the formation of microthrombi. Furthermore, EVs have the potential to activate the coagulation cascade and hinder fibrinolysis, thereby amplifying the risk of thrombosis.\(^7\)\(^4\)

Understanding the proteomic composition of these EVs is crucial for unraveling the molecular mechanisms underlying CAT. MS proteomics was applied to analyze EVs released from

Table 3. Potential biomarkers in cancer-associated thrombosis.

<table>
<thead>
<tr>
<th>Lung cancer</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteins increased in VTE patients</td>
<td>IgV kappa light chain (76)</td>
<td></td>
</tr>
<tr>
<td>Proteins increased in non-VTE patients</td>
<td>Tetranectin (76)</td>
<td></td>
</tr>
<tr>
<td>Non-small cell lung cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proteins increased in VTE patients</td>
<td>SAA1, S100A8, LBP, HP and LDHB (78)</td>
<td></td>
</tr>
<tr>
<td>Pancreas cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proteins increased in VTE patients</td>
<td>IgM Fc, immunoglobulin kappa chain variable region, Ig kappa chainVKIII-JK3, immunoglobulin heavy chain variable region, immunoglobulin kappa light chain variable region (76)</td>
<td></td>
</tr>
<tr>
<td>Proteins increased in non-VTE patients</td>
<td>Immunoglobulin kappa light chain variable region of different sequence (MW 8 kDa), phospholipase D (76)</td>
<td></td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proteins associated with increased risk of cancer associated thrombosis</td>
<td>Angiotensinogen, apolipoprotein B100, CD5 antigen-like, and immunoglobulin heavy constant mu (85)</td>
<td></td>
</tr>
<tr>
<td>Platelet proteome in cancer patients</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uregulated proteins</td>
<td>Lung cancer: Accelerated F13A1, Endoplasmic reticulum proteins (CALR, HSPA5, P4HB) (79)</td>
<td></td>
</tr>
<tr>
<td>Downregulated proteins</td>
<td>Patients with cancer vs healthy controls:FXIII, CALR (82)</td>
<td></td>
</tr>
</tbody>
</table>

CALR, calreticulin; F13A1, factor XII 55 kDa fragment; FXIII, coagulation factor XIII; HP, haptoglobin; HSPA5, heat shock protein family A; LBP, lipopolysaccharide binding protein; LDHB, lactate dehydrogenase B; P4HB, prolyl 4-hydroxylyase subunit beta; SAA1, serum amyloid A1; S100A8, calprotectin; VTE, venous thromboembolism.
platelets in various triple-negative breast cancer (TNBC) cell lines. Results revealed that TNBC cell lines induced platelet aggregation, and the subsequent protein profiling of extracellular vesicles released by platelets highlighted their active participation in this process. Notably, uPAR and PDGFRβ were identified as crucial contributors to the complex mechanism of extracellular vesicle-induced platelet aggregation.90

**Comparing proteomic screening techniques in cancer-associated thrombosis**

There is scant literature comparing proteomics screening techniques in CAT. A study compared multiple reaction monitoring (MRM) MS proteomics with conventional assays to evaluate levels of coagulation factors and fibrinolysis-related proteins. LC-MS was used to profile 31 proteins related to coagulation and fibrinolysis in 75 patients (25 with VTE, 25 with cancer and VTE, and 25 with healthy controls). All samples also underwent traditional antibody or activity-based assays. Both methods had a Pearson correlation of 0.77, indicating a good correlation, but MRM MS had a higher sensitivity, multiplicity, and performance.81

**Limitations of proteomics in cancer-associated thrombosis**

Proteomics has been applied in the context of thrombosis, revealing several promising biomarkers. However, despite these findings, markers have not been globally incorporated into clinical practice. The challenge in using these biomarkers can be attributed to several factors, including limited congruence among study outcomes, substantial variations in methods, protein sample preparation, sample types, and study populations. Proteomics studies in CAT exhibit significant heterogeneity, rendering direct comparisons between investigations challenging. Furthermore, the absence of external study validation adds complexity to the interpretation of results. The majority of these studies had a small sample size, impacting the statistical significance and general applicability of the findings. Subsequent analyses are warranted, with an emphasis on achieving greater methodological similarity across studies.

**Conclusions**

Proteomics enables the comprehensive analysis of protein alterations on a large scale, offering valuable insights for the timely diagnosis, accurate risk assessment, and effective treatment of VTE and CAT. The effective application of biomarkers to clinical practice requires the validation of studies using independent diverse cohorts. Artificial intelligence and machine learning methods are currently under investigation and represent promising tools in combination with proteomics for the identification of biomarkers in thrombosis.

**References**


20. Bruinstroop E, Klok FA, Van De Ree MA, et al. Elevated d-