



(REVIEW ARTICLE)



BRIDGING PROTEOMIC DISCOVERY AND CLINICAL PREDICTION: A REVIEW OF TRANSLATIONAL PATHWAYS FROM CANCER BIOMARKER RESEARCH TO PREDICTIVE SURVIVAL MODELS IN THE UNITED STATES

Agnes Duah ^{1,*}, Godson Teye Apaflo ² and Ebenezer Tetteh ³

¹ School of Public Health – Indiana University Bloomington, Indiana, USA.

² Information Technology and Decision Sciences, University of Energy and Natural Resources, Ghana.

³ Department of Statistics and Actuarial Science – University of Ghana, Legon, Ghana.

* Corresponding Author

Magna Scientia Advanced Biology and Pharmacy, 2026, 19(01), 069–079

Publication history: Received on 26 August 2026; revised on 05 October 2026; accepted on 07 October 2026

Article DOI: <https://doi.org/10.30574/msabp.2026.19.1.0073>

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

Cancer proteomics has generated numerous candidate biomarkers and molecular signatures, but the translation of these discoveries into validated predictive survival models remains incomplete. This systematized review examined the translational pathway from cancer proteomic discovery to predictive survival modeling and clinical implementation in the U.S. oncology context. A structured evidence synthesis of 22 studies was conducted, focusing on biomarker discovery and signature development, survival prediction, analytical and clinical validation, multi-omics and artificial-intelligence-enabled modeling, and clinical translation. The reviewed evidence shows substantial progression from biomarker discovery toward proteomic signatures associated with survival and treatment response and increasingly integrated predictive models. However, the evidence remains uneven across the translational pathway. External validation, analytical reproducibility, integration of molecular and clinical variables, population representation, and prospective evaluation are limited, while direct evidence of clinical utility remains insufficient. The findings indicate that associations between proteomic signatures and survival should not be interpreted as evidence of clinically ready predictive models. Progress toward routine implementation will require rigorous analytical and clinical validation, external evaluation across relevant populations and settings, reproducibility across analytical platforms, and evidence that model outputs improve clinical decision-making.

Keywords: Cancer Proteomics; Biomarker Discovery; Survival Prediction; Precision Oncology.

1 INTRODUCTION

Cancer remains a major clinical and public-health challenge in the United States, with substantial variation in disease progression, treatment response, and survival across cancer types and populations. Growing availability of molecular, clinical, and population-level data has created opportunities to improve risk stratification and prognostic assessment beyond conventional clinicopathological variables (Ye et al., 2023). Cancer proteomics has emerged as an approach for characterizing proteins, signaling processes, and molecular phenotypes that may carry clinically meaningful

information about disease biology. Clinical proteomics has progressed from biomarker discovery and characterization of cellular signaling profiles toward applications intended to support individualized patient management (Calvo et al., 2005; Van Eyk & Dunn, 2007).

Protein expression reflects biological processes closer to the functional phenotype of a tumor than genomic information alone, which makes proteomic profiling relevant to biomarker discovery, identification of biological pathways, prediction of treatment response, and survival assessment (Wong et al., 2022). This relevance is evident at the discovery stage: proteomic analyses have identified predictive proteins associated with survival in high-grade serous ovarian cancer (Xie et al., 2017), developed proteomics-derived signatures associated with chemotherapy response and survival in stage II/III colorectal cancer (Ye et al., 2023), and supported biomarker discovery in colorectal liver metastases (Wong et al., 2022). The central challenge is not the discovery of statistically associated proteins alone, but whether molecular signals identified through high-throughput technologies can be translated into reproducible biomarkers and incorporated into models that improve patient-level prediction.

Moving from discovery toward clinical use generally involves additional evidence concerning verification, analytical performance, clinical validation, predictive performance, and the consequences of using the resulting information in clinical decisions; the extent and ordering of these activities can vary with the intended use (Rodland & Zhang, 2025; Guedes et al., 2025). For the purposes of this review, these activities are grouped into four broader translational evidence stages: discovery, analytical validation, clinical validation and predictive modeling, and clinical translation and utility. These stages are used as an analytic framework for organizing heterogeneous evidence rather than as a claim that every proteomic biomarker must progress through four discrete or strictly sequential phases. Early work positioned proteomics as a pathway toward individualized therapy (Calvo et al., 2005), while more recent work has emphasized moving beyond discovery toward demonstrated clinical utility (Rodland & Zhang, 2025), reflected in work that increasingly connects proteomics with other data types and computational methods. Guedes et al. (2025) consider the integration of digital pathology, proteomics, clinical data, and artificial intelligence analytics, and multi-omics approaches that combine molecular information with artificial intelligence are increasingly investigated for biomarker development and clinical translation (Wang et al., 2026). Within the clinical validation and predictive modeling stage specifically, Maity et al. (2020) demonstrated the relevance of protein-expression data for pan-cancer survival prediction through Bayesian data integration and variable selection, and population-level predictive modeling of U.S. cancer mortality, spanning time-series forecasting and spatial autocorrelation analysis, illustrates how statistical modeling can extend molecular-level discovery toward population-scale outcome prediction (Kubuafor et al., 2025a, 2025b). Taken together, this body of work reflects a conceptual transition from biomarker discovery toward multivariable prediction, and it points to a boundary between proteomic biomarker research and clinical prediction that depends increasingly on computational integration.

The movement from proteomic signatures toward predictive survival models represents a further translational step. Integrative, biomarker-driven approaches connect molecular discovery with clinical management in prostate cancer and precision oncology (Ramesh et al., 2026). Predictive performance alone does not establish clinical usefulness, however: a model may show an association between a molecular signature and survival without improving clinical decisions, which raises methodological considerations concerning validation, bias, and equitable prediction (Amevor, 2026; Amevor et al., 2026a).

Although the existing literature addresses important aspects of proteomic biomarker discovery, protein-expression-based survival prediction, multi-omics integration, and clinical translation, these strands of evidence remain fragmented and do not yet establish a coherent progression across the full translational pathway from molecular discovery to clinically useful survival prediction. In particular, the evidence varies considerably in its consideration of analytical validation, external validation, integration of clinical variables, population heterogeneity, and demonstrated clinical utility. This fragmentation limits a clear assessment of the maturity of the evidence base, making it difficult to determine where support is strongest, where critical translational bottlenecks occur, and which stages remain insufficiently supported for reliable clinical application. The central review question is: How does the evidence base support progression from cancer proteomic biomarker discovery to validated predictive survival models and, ultimately, clinically useful application in oncology?

This review therefore examines the translational pathway linking cancer proteomic discovery with predictive survival modeling, with particular attention to the distinction between biomarker discovery, validation, prediction, and clinical utility. It synthesizes evidence across these stages, identifies methodological and evidentiary gaps that limit translation, and examines how proteomic information is integrated with clinical, genomic, computational, and artificial-intelligence approaches within the U.S. cancer research context. The review further considers predictive performance, population heterogeneity, and translational readiness to clarify the distance between promising proteomic signals and clinically defensible predictive applications.

2 METHODS

2.1 Information Sources and Evidence Identification

This review used a structured evidence-identification approach to assemble studies relevant to the pathway from cancer proteomic discovery to predictive survival modeling. The evidence corpus was developed around concepts including proteomics, protein expression, cancer biomarkers, survival prediction, precision oncology, multi-omics, artificial intelligence, and clinical translation. Because the contemporaneous database-specific search records, search dates, and complete search syntax used to construct the corpus were not retained, the precise retrieval strategy cannot be reproduced in full. Accordingly, no specific databases, search dates, or search strings are retrospectively attributed to the study. The resulting records were screened against the eligibility criteria described in Section 2.2.

The initial evidence pool comprised **238 potentially relevant records**. Duplicate records were removed before relevance assessment, followed by screening for cancer or cancer-biomarker relevance, relevant molecular, prognostic, or predictive components, contribution to the proposed translational pathway, and methodological relevance to the review question. This staged process resulted in a final evidence corpus of **22 studies**. The resulting evidence was synthesized narratively to provide a focused assessment of the progression from proteomic discovery through validation and predictive modeling to clinical translation and utility.

2.2 Conceptual Eligibility Framework: Inclusion and Exclusion Criteria

A modified PICOS framework was used to determine whether each reference was sufficiently relevant to the review question. Because the objective was to examine a translational pathway rather than a single intervention, the framework incorporated biomarker discovery, predictive modeling, and clinical-translation evidence. Table 1 below summarizes the evidence selection criteria for the study.

Table 1: Eligibility criteria applied to the evidence corpus

Criterion	Inclusion	Exclusion
Population/context	Human cancer, oncology, cancer biomarker, cancer survival, mortality, or precision-oncology contexts	Studies unrelated to cancer or oncology
Molecular evidence	Proteomics, protein expression, clinical proteomics, proteomic signatures, or integrated molecular data involving proteomic information	Studies with no meaningful molecular/biomarker relevance
Clinical outcome	Survival, mortality, treatment response, prognosis, risk prediction, early detection, or clinically relevant cancer outcomes	Studies without a meaningful cancer-related outcome or translational endpoint
Prediction/modeling	Survival models, prognostic models, predictive signatures, AI/ML models, statistical prediction, or computational integration	Purely descriptive studies with no relevant connection to biomarker development or prediction
Translational relevance	Evidence addressing biomarker discovery, verification, validation, clinical prediction, precision oncology, clinical utility, or implementation	Research with no plausible connection to the discovery-to-clinical pathway
Geographic relevance	U.S.-specific evidence was prioritized; international evidence was retained where it informed the general translational pathway	Evidence whose subject matter had no meaningful relevance to the review's translational question
Publication type	Peer-reviewed research, reviews, perspectives, methodological studies, and other scholarly sources included in the supplied corpus	Non-scholarly material, duplicated references, or material clearly outside the research question

This approach was necessary because the 22-study corpus contains different levels of translational evidence. Some studies provide direct evidence connecting proteomic information with survival-related outcomes, others address statistical prediction using protein-expression data, and others address broader clinical integration, artificial intelligence, precision oncology, or clinical utility. Table 1 identifies which category each study falls into.

2.3 Study Exclusion Logic

Studies were excluded from the core evidence synthesis if they fell outside the scope of cancer or cancer-related biomarker research, contained no relevant proteomic, molecular, prognostic, or predictive component, or did not contribute to any stage of the proposed translational pathway. Duplicate studies were also excluded, as were studies whose methodological content could not reasonably inform the translation from cancer biomarker discovery to predictive application. Studies that were not directly proteomic in focus but that contributed to the predictive or translational stages of the pathway, such as those addressing gene expression, artificial intelligence, or population-level mortality, were retained as contextual evidence rather than treated as direct proteomic investigations. This distinction was maintained throughout the review to avoid misrepresenting non-proteomic studies as proteomic research, and to preserve the accuracy of the evidence base underlying the translational pathway under examination.

2.4 Study Selection and Relevance Assessment

An initial search identified 238 studies as potentially relevant to the review's central research question. Following the removal of 34 duplicate records, 204 studies proceeded to relevance screening. At this stage, 61 studies were excluded for falling outside the scope of cancer or cancer-related biomarker research, leaving 143 studies for further assessment. A subsequent screening step removed 58 studies that contained no relevant proteomic, molecular, prognostic, or predictive component, reducing the pool to 85 studies. Of these, 47 were excluded for not contributing to any stage of the proposed translational pathway from proteomic discovery to survival prediction, leaving 38 studies. A final assessment excluded 16 studies whose methodological content could not reasonably inform the translation from cancer biomarker discovery to predictive application. This staged process of duplicate removal and criterion-based exclusion resulted in a final corpus of 22 studies retained for inclusion in the review.

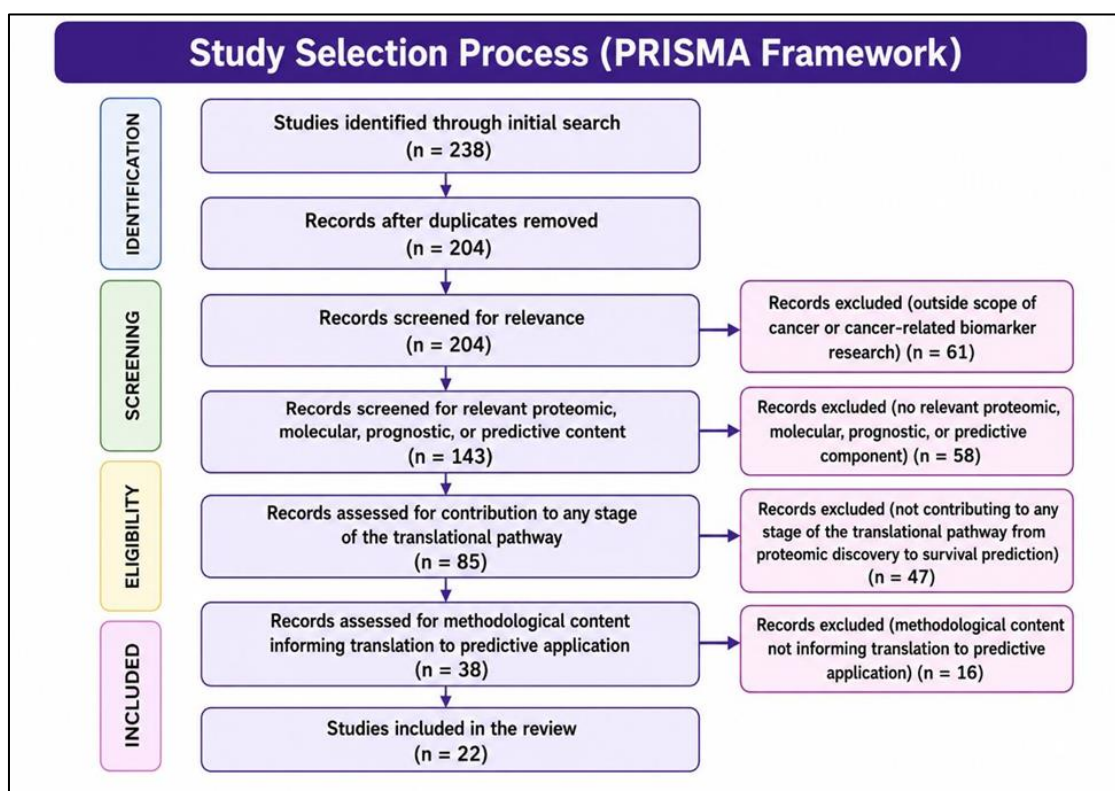


Figure 1: PRISMA Framework for the Study Selection Process

Source: Page et al. (2021)

2.5 Data Extraction and Methodological Limitations of the Evidence Base

A structured extraction framework was used to organize information from the 22 studies, capturing the author and publication year, cancer type or clinical context, and study objective for each. Additional information extracted included the molecular or proteomic data type employed, the biomarker or molecular signature under investigation, and the analytical or computational approach used to generate findings. Each study was further characterized in terms of its clinical outcome and survival or prognostic relevance, its integration with clinical or other molecular data, and its predictive or translational objective. Finally, each study's contribution to the proteomic-to-clinical-prediction pathway

was recorded, allowing the extracted data to be positioned within the broader translational framework guiding the review.

2.6 Methodological Limitations

The evidence base presents several limitations that were considered during synthesis. The 22 studies are methodologically heterogeneous, spanning original research articles, reviews, perspectives, a book, and preprints, and cannot be treated as a uniform set of clinical validation studies. They also address different stages of translation: clinical proteomics literature focuses on biomarker discovery and individualized therapy, while other studies focus on protein-expression-based survival prediction, AI-enabled multi-omics, or clinical utility. Cancer types and outcomes differ substantially across the corpus, including ovarian, colorectal, prostate, colorectal liver metastases, melanoma, NSCLC, and population-level cancer mortality, limiting direct quantitative comparison. Again, U.S. representation is uneven, with some studies examining U.S. cancer mortality and lung cancer mortality prediction specifically, while several proteomic studies draw on other populations or provide general translational evidence. Finally, predictive-performance metrics are not reported consistently across studies, so the evidence was not pooled into a meta-analysis. Instead, the review uses a structured narrative synthesis centered on the translational stages connecting biomarker discovery with clinical prediction.

2.7 Synthesis Strategy

Given the heterogeneity of the evidence, a stage-based narrative synthesis was used rather than statistical meta-analysis. Accordingly, this review is best characterized as a systematized review (Grant & Booth, 2009): it applies a systematic, staged search and screening process, reported using PRISMA-style flow documentation, but relies on narrative rather than meta-analytic synthesis and does not apply a named per-study quality-appraisal instrument such as AMSTAR, ROBIS, or GRADE. The synthesis was organized around four translational evidence stages: discovery, analytical validation, clinical validation and predictive modeling, and clinical translation and utility. Each study was assigned a principal position within this pathway, while recognizing that individual studies could contribute to more than one stage. The synthesis distinguished association from prediction throughout: a protein associated with survival was not automatically treated as a validated predictive biomarker, and a statistically accurate prediction model was not treated as evidence of clinical usefulness. The translational pathway therefore progressed from the identification of cancer-associated proteins and proteomic profiles, through analytical and biological verification of candidate biomarkers, to clinical validation and predictive survival modeling, and ultimately to assessment of clinical utility and implementation potential. This framework emphasizes that successful translation requires evidence to progress beyond biomarker discovery toward reproducible validation, clinically meaningful prediction, and demonstrable utility in decision-making. This approach provides the basis for the subsequent Results of the Search, where the 22-study corpus is presented and organized according to the four translational evidence stages.

3 FINDINGS OF THE STUDY

3.1 Study Characteristics

The 22 studies included in the review span original research, reviews, perspectives, a book, and preprints, consistent with the heterogeneity described in the Methods. Table 2 presents the full evidence corpus and identifies each study's principal position within the four-stage translational framework. Studies that provide contextual rather than direct proteomic evidence are explicitly identified as such.

Table 2: Distribution of the Included Studies Across the Four-Stage Translational Pathway

Study	Cancer type / context	Proteomic, predictive, or molecular, analytical approach	Principal evidence contribution / translational position
Calvo et al. (2005)	General oncology	Clinical proteomics, biomarker discovery, cell-signaling profiles	Discovery - establishes clinical proteomics as a pathway from biomarker discovery toward individualized cancer therapy
Van Eyk and Dunn (2007)	General oncology	Clinical proteomics, protein profiling and diagnostic/therapeutic applications	Discovery - provides foundational clinical proteomics framework linking protein measurement with cancer diagnosis and treatment

Xie et al. (2017)	High-grade serous ovarian cancer	Proteomic profiling, pathway analysis and identification of predictive proteins	Discovery - identifies proteins and biological pathways associated with survival
Barillot et al. (2020)	Cancer systems biology	Computational systems biology, molecular networks and integrative cancer analysis	Discovery - provides computational framework for connecting molecular information with cancer biology and downstream prediction
Maity et al. (2020)	Pan-cancer	Protein-expression data integration, Bayesian variable selection and survival prediction	Clinical validation and predictive modeling - directly demonstrates protein-expression-based survival prediction
Wong et al. (2022)	Colorectal liver metastases	Proteomic profiling and biomarker discovery	Discovery - identifies proteomic features and candidate biomarkers relevant to metastatic colorectal cancer
Ye et al. (2023)	Stage II/III colorectal cancer	High-throughput proteomic profiling and prognostic signature development	Clinical validation and predictive modeling - links a proteomic signature with chemotherapy response and survival
Aqel et al. (2023)	Untreated advanced melanoma	Economic evaluation of relatlimab plus nivolumab versus nivolumab	Clinical translation and utility - contextual - provides evidence concerning treatment value and economic implications of cancer outcome-related decision-making
Rubinstein et al. (2024)	Multicancer detection	Translational evaluation of multicancer detection testing and clinical implementation	Clinical translation and utility - addresses movement of molecular cancer detection technologies toward clinical application
Aliu et al. (2025)	Advanced NSCLC with KRAS G12C mutation	Cost-effectiveness and value-of-information analysis	Clinical translation and utility - contextual - informs treatment decision-making, economic value and uncertainty surrounding targeted cancer therapy
Guedes et al. (2025)	General oncology	Integration of digital pathology, proteomics, clinical data and AI analytics	Clinical translation and utility - addresses integration of proteomic and clinical information with AI toward translational cancer research
Kubuafor et al. (2025a)	U.S. lung cancer mortality	Time-series forecasting of lung cancer mortality rates	Clinical validation and predictive modeling - contextual - demonstrates application of predictive modeling to U.S. cancer outcomes, without direct proteomic analysis
Kubuafor et al. (2025b)	U.S. cancer mortality	Spatial clustering, spatial heterogeneity and spatial autocorrelation	Clinical validation and predictive modeling - contextual - contributes population-level analytical evidence relevant to cancer outcome prediction
Rodland and Zhang (2025)	General oncology	Comprehensive cancer proteomic datasets and translational clinical insight	Clinical translation and utility - focuses on converting large-scale cancer proteomic datasets into clinically meaningful information
Amevor (2026)	Cross-cancer	AI-biostatistical frameworks, bias mitigation and cancer outcome prediction	Clinical translation and utility - addresses equitable predictive analytics and implementation considerations for cancer outcome models
Amevor et al. (2026a)	Cancer prediction / early cancer detection	AI-based predictive models, statistical considerations and validation methods	Clinical validation and predictive modeling - addresses validation requirements and statistical

			considerations for AI-enabled cancer prediction
Amevor et al. (2026b)	Real-world NSCLC	Time-varying hazard patterns and KRAS G12C/G12D co-mutation profiles	Clinical validation and predictive modeling - contextual - examines molecular and survival-related patterns relevant to outcome prediction
Ramesh et al. (2026)	Prostate cancer	Integrative biomarker-driven management, genomic discovery and precision oncology	Clinical translation and utility - examines progression from molecular biomarker discovery toward precision cancer management
Therachiyil et al. (2026)	Cancer, particularly early detection and targeted therapy	Proteomics-driven biomarker discovery and therapeutic applications	Clinical translation and utility - connects proteomic biomarkers with early detection and targeted treatment applications
Wang et al. (2026)	Cancer immunotherapy / immune checkpoint blockade	AI-enabled multi-omics biomarkers, predictive modeling and clinical translation	Clinical translation and utility - integrates multi-omics and AI for predictive biomarker development and treatment-response assessment
Liu et al. (2025)	Cancer / biomarker research	Cancer biomarker discovery and translation from discovery toward treatment	Clinical translation and utility - explicitly addresses the transition from biomarker discovery to treatment-oriented application
Barnabas et al. (2019)	Ovarian cancer	Proteomic analysis of uterine lavage fluid as a proximal liquid biopsy	Discovery - identifies a candidate proteomic biomarker panel from uterine lavage fluid, positioned as an alternative to CA125 and ultrasound screening, aiming for earlier detection with greater specificity

3.2 Design-Level Bias Considerations

Table 3: Design-level bias considerations across the four translational phases

Phase	Predominant design	U.S. population representation	Principal bias consideration
Discovery	Discovery cohort, laboratory-based proteomic screening	Not consistently specified	Small discovery cohorts limit generalizability
Analytical validation	Not well represented in the corpus	Not specified	Absence of reported inter-laboratory or platform standardization data
Clinical validation and predictive modeling	Retrospective cohort, population-level modeling	Partial (Kubuafor et al., 2025a and Amevor et al., 2026b explicitly U.S.-based; Kubuafor et al., 2025b also U.S.-based)	Cancer-type and outcome heterogeneity limits pooled comparison
Clinical translation and utility	Review, perspective, economic evaluation	Mixed, largely general or non-U.S.-specific	Predictive-performance metrics not consistently reported

Sources: (Kubuafor et al., 2025a, 2025b; Wang et al., 2026; Amevor et al., 2026b)

Consistent with the systematized-review approach described in Section 2.5, no formal per-study quality-appraisal instrument (e.g., AMSTAR, ROBIS, GRADE) was applied. Table 3 instead summarizes design-level bias considerations across the four translational phases, based on study design type and the extent to which U.S. populations are represented.

The 22 studies were classified according to their principal contribution to the four-stage translational pathway. Six (6) references primarily addressed proteomic biomarker discovery, no reference (0) primarily addressed analytical validation, six (6) references addressed clinical validation and predictive modeling, and ten (10) references addressed clinical translation and utility. One additional study, Page et al. (2021), provided the methodological framework for reporting systematic reviews and was therefore not assigned to a substantive translational stage. Thus, the four substantive components accounted for all 22 studies in Table 2, with Page et al. (2021) providing the review's methodological framework separately. The findings therefore indicate that the evidence base is concentrated in proteomic biomarker discovery, clinical validation and predictive modeling, and clinical translation and utility, while analytical validation is not directly represented as a principal focus.

4 DISCUSSION AND TRANSLATIONAL SYNTHESIS

4.1 Synthesis of Translational Pathways

The evidence reveals an uneven progression from proteomic discovery toward clinically useful survival prediction. Across the corpus, discovery studies consistently demonstrate that protein-level measurements can identify cancer-associated signals and candidate biomarkers, but the evidence becomes thinner as the pathway moves toward reproducible validation and demonstrated clinical utility. This pattern is more informative than the number of studies assigned to each phase, because the studies differ substantially in cancer type, data source, analytical method, and translational objective.

- **Phase I:** Discovery to verification. Early clinical proteomics work established the rationale for using protein measurements to characterize tumor biology and identify candidate biomarkers (Calvo et al., 2005; Van Eyk & Dunn, 2007). Xie et al. (2017) extended this approach by identifying proteins associated with survival in high-grade serous ovarian cancer, while Wong et al. (2022) used proteomic profiling to investigate biomarker discovery in colorectal liver metastases. These studies converge in showing that proteomic measurements can generate biologically and clinically relevant candidate signals, but they also illustrate the dependence of discovery findings on cancer type, biological context, and study design. Their contribution is strongest at the signal-generation stage rather than at establishing reproducible clinical biomarkers.
- **Phase II:** Analytical validation. The transition from candidate discovery to analytically reliable measurement is poorly represented in the corpus. The reviewed studies provide limited direct evidence concerning reproducibility, sensitivity, cross-platform consistency, or inter-laboratory standardization. This contrasts with the comparatively richer evidence for biomarker discovery and survival modeling. No study in the corpus explicitly reports CLIA or CAP compliance data for a proteomic assay, further limiting what can be concluded about readiness for routine clinical testing.
- **Phase III:** Clinical validation and predictive modeling. The corpus contains more modeling-oriented evidence than direct evidence addressing analytical validation, although the studies differ substantially in what they establish. Ye et al. (2023) evaluated a proteomics-derived signature associated with chemotherapy response and survival in stage II/III colorectal cancer, whereas Maity et al. (2020) developed a Bayesian survival-prediction approach using protein-expression data across cancers. These studies demonstrate that molecular information can contribute to prognostic or predictive modeling, but they should not be treated as equivalent to prospective clinical validation. The Kubuafor et al. (2025a, 2025b) and Aमेvor et al. (2026b) studies add U.S. population-level evidence on lung cancer mortality forecasting, time-varying hazard patterns and co-mutation profiles in real-world NSCLC, and spatial patterns in cancer mortality; their role in this corpus is contextual rather than direct validation of a proteomic signature. Considered together, the studies in this phase show that the reviewed corpus contains more evidence addressing predictive modeling than direct evidence addressing analytical validation, while also showing that model performance, prognostic association, and population-level forecasting do not by themselves establish clinical readiness.
- **Phase IV:** Clinical translation and utility. The largest share of the corpus addresses translation conceptually rather than through direct implementation studies. Guedes et al. (2025) consider integration of digital pathology, proteomics, clinical data, and AI analytics, while Wang et al. (2026) examine multi-omics and AI-enabled prediction. Ramesh et al. (2026) situate biomarker-driven approaches within prostate cancer precision oncology, and Rodland and Zhang (2025) and Therachiyil et al. (2026) discuss the translational potential of proteomic datasets. These studies help define how proteomic information might be incorporated into clinically oriented systems, but they do not demonstrate that a proteomic survival model improves patient-level decisions or outcomes. The economic evaluations reported by Aliu et al. (2025) and Aqel et al. (2023) provide contextual evidence that economic evaluation can be applied to molecularly defined cancer treatment strategies. Still, they are not direct evaluations of proteomic prognostic assays. The corpus therefore provides substantial evidence concerning clinical translation and potential utility, but limited direct evidence of demonstrated clinical utility for proteomic survival models. Within the reviewed corpus, there is no reported decision-curve analysis, prospective implementation study, documented electronic-health-record integration,

or direct demonstration that use of a proteomic survival model changes patient-level decisions or outcomes. These forms of evidence can help establish clinical utility, but the appropriate evaluation depends on the intended clinical decision, setting, and consequences of model use. Accordingly, clinical utility remains an important translational objective rather than an established outcome in the reviewed corpus. Taken together, the evidence suggests that the principal translational bottleneck is not biomarker discovery alone. The more difficult transition occurs between identifying survival-relevant molecular signals, demonstrating their reproducibility and generalizability, and establishing that their use changes clinical decisions in a beneficial way. The studies therefore support a progression in technical capability without a matching progression in clinical validation or real-world impact. This distinction explains why the pathway from proteomic discovery to routine clinical use remains incomplete despite substantial progress in molecular profiling and predictive modeling.

4.1.1 Evidence Across the Translational Pathway

Comparison across the four translational phases reveals that the evidence base changes not only in quantity but also in evidentiary function as studies progress toward clinical application. Discovery studies consistently identify biologically relevant proteins and candidate biomarkers across multiple cancer types, providing strong evidence for signal identification but relatively little evidence for reproducibility beyond the original study context. Analytical-validation evidence is notably sparse, creating the largest discontinuity in the translational pathway because few studies evaluate cross-platform consistency, inter-laboratory reproducibility, or assay standardization.

The literature becomes substantially richer during the clinical-validation-and-predictive-modeling phase, where studies increasingly integrate proteomic information with survival outcomes, Bayesian prediction, protein-expression data, or population-level modeling. However, these studies differ considerably in their analytical objectives and evidentiary strength. Some develop prognostic models using molecular information, whereas others contribute contextual evidence about mortality prediction or population heterogeneity rather than validating proteomic assays directly. This explains why predictive capability appears to advance more rapidly than analytical validation within the reviewed corpus.

The clinical-translation-and-utility phase contains the largest number of studies, yet most address translation conceptually rather than empirically. Reviews, perspectives, multi-omics frameworks, artificial-intelligence integration, and economic evaluations collectively describe how proteomic information could be incorporated into clinical systems, but relatively few studies demonstrate prospective implementation, decision-curve analysis, electronic-health-record integration, or patient-level improvements resulting from proteomic survival models. Considered together, the corpus indicates that the principal translational bottleneck occurs between predictive modeling and demonstrated clinical utility rather than between biomarker discovery and model development.

4.2 Barriers to Clinical Implementation in the U.S.

Demographic representation is an important consideration, but the corpus provides stronger evidence of population heterogeneity than of proteomic-model transportability. Work on equitable AI-biostatistical frameworks identifies bias mitigation as an important consideration for predictive models developed from non-representative cohorts (Amevor, 2026; Amevor et al., 2026a), while the U.S. mortality studies demonstrate variation in cancer outcomes across geographic and population settings (Kubuafor et al., 2025a; Amevor et al., 2026b). These findings indicate that external validation across relevant populations should be treated as an important requirement for proteomic survival models. However, the reviewed corpus does not provide sufficient direct evidence to determine whether specific proteomic signatures maintain discrimination, calibration, or overall predictive performance across demographic or geographic populations. Generalizability therefore remains an unresolved validation question rather than a demonstrated source of model failure.

4.3 Future Research Directions

The corpus points toward three directions for closing the discovery-to-prediction gap. Multi-omics integration, combining proteomic measurement with genomic data and clinical artificial intelligence, is already the stated focus of several recent studies in the corpus (Wang et al., 2026; Guedes et al., 2025), suggesting this is the direction most actively pursued rather than merely proposed. Decentralized testing, moving proteomic assays from centralized reference laboratories toward point-of-care or local hospital networks, is not directly evidenced within the reviewed corpus and represents an infrastructure question the current literature has not yet addressed empirically. Prospective, biomarker-driven adaptive trial designs are similarly underrepresented: the corpus contains retrospective and modeling studies but no prospective trial reported around a proteomic signature specifically, which is consistent with the broader finding that Phase III and Phase IV evidence in this review remains largely associational and conceptual rather than confirmatory.

5 CONCLUSION

Cancer proteomic research has progressed substantially from biomarker discovery toward prognostic and predictive applications, but the translation of molecular signals into clinically reliable survival models remains uneven. The principal bottlenecks are not limited to biomarker identification. They include analytical and clinical validation, integration of molecular and clinical information, reproducibility across datasets and analytical platforms, external validation across relevant populations and settings, and demonstration of meaningful clinical utility. Across the corpus, evidence of survival association represents an earlier translational stage than evidence that a proteomic model can reliably guide patient-level decisions. Routine clinical implementation will require a more complete progression from discovery and analytical validation through clinical validation and predictive modeling to clinical translation, with evidence of clinical utility appropriate to the intended clinical use. Multi-omics, imaging, clinical data, and artificial intelligence may strengthen predictive systems, but technical integration alone will not establish clinical value. Future work should therefore place greater emphasis on external validation, prospective evaluation, reproducibility, and evidence that model outputs improve clinical decision-making. The translational pathway from proteomic discovery to routine cancer care is therefore promising, but its clinical maturation depends on closing the gap between predictive capability and demonstrated patient-level benefit.

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Disclosure of conflict of interest

The authors declare no conflicts of interest regarding this manuscript.

REFERENCES

- [1] Aliu, O. B., Erstad, B. L., Liang, J., & Abraham, I. (2025). EE79 cost-effectiveness and value of information analyses of sotorasib versus docetaxel in previously treated patients with advanced non-small cell lung cancer with KRAS G12C mutation. *Value in Health*, 28(6), S76. <https://doi.org/10.1016/j.jval.2025.04.371>.
- [2] Aqel, O., Alfayoumi, I., Shen, Y., Aliu, O., Zaidi, S. H., Moreno-Rodriguez, V., Ayoade, O. O., Alzahir, H., Kopecky, J. M., Khattab, S., Snyder, M., & Abraham, I. (2023). Economic evaluation for the US of relatlimab and nivolumab (RELA+NIVO) versus nivolumab (NIVO) monotherapy in untreated advanced melanoma. *Journal of Clinical Oncology*, 41(16_suppl), 9539. https://doi.org/10.1200/JCO.2023.41.16_suppl.9539.
- [3] Amevor, R. (2026). Equitable AI-biostatistical frameworks for bias mitigation in cancer prediction and outcome analytics: A narrative review. *EPRA International Journal of Research & Development (IJRD)*, 479. <https://doi.org/10.36713/epra26770>.
- [4] Amevor, R., Aduampong, M. J. K., & Nketiah, T. A. (2026a). Evaluation of AI-based predictive models for early cancer detection: Statistical considerations and validation methods. *Journal of Innovative Science*, 2(1), 1–10.
- [5] Amevor, R., Baidoo, D., & Kubuafor, E. (2026b). Time-varying hazard patterns and co-mutation profiles of KRAS G12C and G12D in real-world NSCLC. *arXiv*. <https://arxiv.org/abs/2602.19295>
- [6] Barnabas, G. D., Bahar-Shany, K., Sapoznik, S., Helpman, L., Kadan, Y., Beiner, M., Weitzner, O., Arbib, N., Korach, J., Perri, T., Katz, G., Blecher, A., Brandt, B., Friedman, E., Stockheim, D., Jakobson-Setton, A., Eitan, R., Armon, S., Brand, H., Zadok, O., Aviell-Ronen, S., Harel, M., Geiger, T., & Levanon, K. (2019). Microvesicle proteomic profiling of uterine liquid biopsy for ovarian cancer early detection. *Molecular & Cellular Proteomics*, 18(5), 865–875. <https://doi.org/10.1074/mcp.RA119.001362>
- [7] Barillot, E., Calzone, L., Hupé, P., Vert, J. P., & Zinovyev, A. (2020). *Computational systems biology of cancer*. CRC Press.
- [8] Calvo, K. R., Liotta, L. A., & Petricoin, E. F. (2005). Clinical proteomics: From biomarker discovery and cell signaling profiles to individualized personal therapy. *Bioscience Reports*, 25(1-2), 107–125. <https://doi.org/10.1007/s10540-005-2851-3>
- [9] Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848>
- [10] Guedes, J., Woldmar, N., Szasz, A. M., Wieslander, E., Pawlowski, K., Horvatovich, P., Malm, J., Szadai, L., Nemeth, I. B., Marko-Varga, G., & Gil, J. (2025). A perspective on integrating digital pathology, proteomics, clinical data and AI analytics in cancer research. *Journal of Proteomics*, 320, Article 105493. <https://doi.org/10.1016/j.jprot.2025.105493>

- [11] Kubuafor, E., Baidoo, D., Okeke, O. J., Amevor, R., Arhin, G., & Korley, J. T. (2025a). Evaluation of time series forecasting models for predicting lung cancer mortality rates in the United States: A comparison with Altuhaifa (2023) study. arXiv. <https://arxiv.org/abs/2508.16052>
- [12] Kubuafor, E., Baidoo, D., Duah, A., Amevor, R., Okeke, O. J., Quaye, D., & Appiah, P. O. (2025b). Latent spatial heterogeneity in U.S. cancer mortality: A multi-site clustering and spatial autocorrelation analysis. arXiv. <https://arxiv.org/abs/2508.12188>
- [13] Liu, H., Karsidag, I., Golin, R., & Wu, G. (2025). Bridging Discovery and Treatment: Cancer Biomarker. *Cancers*, 17(22), 3720. <https://doi.org/10.3390/cancers17223720>
- [14] Maity, A., Bhattacharya, A., Mallick, B. K., & Baladandayuthapani, V. (2020). Bayesian data integration and variable selection for pan-cancer survival prediction using protein expression data. *Biometrics*, 76(1), 316–325. <https://doi.org/10.1111/biom.13132>
- [15] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- [16] Ramesh, P., Shree, L., & Balraj, A. S. (2026). Integrative advances in biomarker-driven prostate cancer management from genomic discovery to precision oncology. *Discover Oncology*, 17, 1060. <https://doi.org/10.1007/s12672-026-05311-7>
- [17] Rodland, K. D., & Zhang, B. (2025). Toward real clinical utility: Leveraging comprehensive cancer proteomic datasets for clinical insight. *Expert Review of Proteomics*, 22(10), 375–387. <https://doi.org/10.1080/14789450.2025.2580647>
- [18] Rubinstein, W. S., Patriotis, C., Dickherber, A., Han, P. K. J., Katki, H. A., LeeVan, E., Pinsky, P. F., Prorok, P. C., Skarlupka, A. L., Temkin, S. M., Castle, P. E., & Minasian, L. M. (2024). Cancer screening with multicancer detection tests: A translational science review. *CA: A Cancer Journal for Clinicians*, 74(4), 368–382. <https://doi.org/10.3322/caac.21833>
- [19] Therachiyil, L., Surendranath, A., Anand, A., Dogan Sahin, R., Al-Shabeeb Akil, A. S., Bhat, A. A., & Uddin, S. (2026). Proteomics-driven cancer biomarkers for early detection and targeted therapy: Insights from the Middle East. *Journal of Proteome Research*, 25(1), 21–40. <https://doi.org/10.1021/acs.jproteome.5c00949>
- [20] Van Eyk, J. E., & Dunn, M. J. (Eds.). (2007). *Clinical proteomics: From diagnosis to therapy*. Wiley-VCH. <https://doi.org/10.1002/9783527622153>
- [21] Wang, X., Xiong, D., Cui, S., Duan, B., Ding, G., Huang, Y., & Wang, Q. (2026). Artificial intelligence-enabled multi-omics biomarkers for immune checkpoint blockade: Mechanisms, predictive modeling, and clinical translation. *Frontiers in Immunology*, 17, 1732079. <https://doi.org/10.3389/fimmu.2026.1732079>
- [22] Wong, G. Y. M., Diakos, C. I., Hugh, T. J., & Molloy, M. P. (2022). Proteomic profiling and biomarker discovery in colorectal liver metastases. *International Journal of Molecular Sciences*, 23(11), 6091. <https://doi.org/10.3390/ijms23116091>
- [23] Xie, H., Wang, W., Sun, F., Deng, K., Lu, X., Liu, H., Zhao, W., Zhang, Y., Zhou, X., Li, K., & Hou, Y. (2017). Proteomics analysis to reveal biological pathways and predictive proteins in the survival of high-grade serous ovarian cancer. *Scientific Reports*, 7, 9896. <https://doi.org/10.1038/s41598-017-10559-9>
- [24] Ye, S.-B., Cheng, Y.-K., Li, P.-S., Zhang, L., Zhang, L.-H., Huang, Y., Chen, P., Wang, Y., Wang, C., Peng, J.-H., Shi, L.-S., Ling, L., Wu, X.-J., Qin, J., Yang, Z.-H., & Lan, P. (2023). High-throughput proteomics profiling-derived signature associated with chemotherapy response and survival for stage II/III colorectal cancer. *npj Precision Oncology*, 7, 50. <https://doi.org/10.1038/s41698-023-00400-0>