Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Molecular Biomarkers for Predicting Platinum Resistance in Ovarian Cancer

Fatima Abdulwahab

abdulwahabfatima43@gmail.com

Bridgewater State University

Research Article

Keywords: Keywords: ovarian cancer; platinum resistance; molecular biomarkers; external validation; gene expression; homologous recombination

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11849967/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Molecular Biomarkers for Predicting Platinum Resistance in Ovarian Cancer

An external assessment of seven candidate transcripts in two public microarray cohorts

Fatima Abdulwahab

Independent Researcher, Dallas, Texas, USA

ORCID: 0009-0008-2298-8386 · Correspondence: abdulwahabfatima43@gmail.com

Abstract

Molecular mechanisms of platinum resistance in ovarian cancer are well described, but translating mechanistic candidates into reproducible pretreatment predictors remains difficult. This study examined the portability of seven literature-selected transcripts using two public microarray cohorts. GSE51373 contained 28 high-grade serous tumours classified by markedly different progression-free survival intervals after first-line chemotherapy. The primary external assessment used 70 serous tumours with deposited platinum-sensitivity labels from GSE63885, including 32 resistant and 38 sensitive cases. *ABCB1*, *BRCA1*, *BRCA2*, *CCNE1*, *ERCC1*, *RAD51* and *SLFN11* were summarised using outcome-independent probe mapping. Each score direction was selected in GSE51373 and then fixed before external assessment; *ERCC1* showed an exact discovery tie and was oriented by a prespecified tie rule. Discovery areas under the receiver operating characteristic curve ranged from 0.500 to 0.667. In the external serous cohort, oriented areas ranged from 0.434 to 0.615. *ABCB1* had the highest external area at 0.615, with a stratified bootstrap 95% interval of 0.485–0.744. No gene met a false discovery rate threshold of 0.05 in either cohort. Expanding the external cohort to all 75 labelled tumours produced similar weak discrimination. These results illustrate limited portability of these transcript summaries under heterogeneous outcome definitions and preprocessing. They support further study of assay-specific and functional measurements with harmonised clinical endpoints, rather than clinical use of the evaluated expression scores.

Keywords: ovarian cancer; platinum resistance; molecular biomarkers; external validation; gene expression; homologous recombination

1 Introduction

Platinum-based chemotherapy remains central to the treatment of many epithelial ovarian cancers. Resistance can be present at initial treatment or emerge following an earlier response, and its clinical interpretation depends on treatment history and the interval to recurrence. Ovarian cancer also comprises distinct histological diseases with different molecular features and response patterns [13, 14]. A useful pretreatment biomarker must therefore do more than identify a plausible resistance mechanism. It must produce a reproducible measurement in a specified population, predict a defined outcome and add information relevant to a treatment decision.

DNA repair provides a strong biological rationale for biomarker development. Germline and somatic alterations in homologous recombination genes have been associated with platinum

response and survival [22]. Conversely, restoration of BRCA1 or BRCA2 function through secondary alterations can accompany acquired resistance [20, 21, 26]. Other candidate pathways include replication stress, drug transport and damage tolerance. However, measurements at different biological levels can give different information.

BRCA1 transcript abundance cannot establish whether a pathogenic *BRCA1* variant is present. *RAD51* mRNA cannot show whether RAD51 is being recruited into nuclear repair foci. A genomic scar records a history of defective repair, whereas a functional assay interrogates repair activity in the sampled material [17, 23, 28].

Much of the practical difficulty lies in moving between these levels of evidence. Studies can identify a statistically associated transcript in a small selected cohort and then interpret it through a mechanism established in cell models or through a clinically used assay with a different readout. A response-associated signal can also reflect tumour grade, stromal composition, treatment regimen or other determinants of recurrence. An external assessment that uses the same expression summary and a fixed score direction provides a limited but direct test of whether the ranking survives a change of cohort.

This study assesses seven biologically motivated transcripts in two archived cohorts. It asks whether the direction chosen in a small discovery set provides useful discrimination in a separate serous cohort and whether the result changes when the external cohort is expanded beyond serous histology. The purpose is to evaluate portability and measurement limits. The study develops no multigene classifier and identifies no treatment-selection threshold. Its contribution is a reproducible candidate-marker reanalysis with explicit sample accounting, fixed external scoring and uncertainty reporting.

2 Biological evidence and the predictive question

2.1 Response definitions and treatment context

Clinical platinum responsiveness is related to, but incompletely captured by, time to progression. Current guidance considers prior response, treatment-free interval, tolerability and other clinical circumstances when assessing whether platinum is appropriate [8]. Historical datasets often simplify that judgement into categorical labels. In one dataset used here, the groups were selected using progression-free survival below eight months and above eighteen months. In the other, the deposited sensitivity categories refer to disease-free survival thresholds. These outcomes include the effects of surgery, baseline disease and follow-up, and they are not identical to radiological response after a fixed number of cycles.

A marker associated with outcome among patients receiving one regimen can be useful as a response-associated or prognostic marker. Demonstrating that it predicts differential benefit from platinum requires a treatment comparison and an interaction analysis or another design capable of addressing that question. Both cohorts here received platinum-containing treatment. Consequently, the title's predictive question is approached through discrimination of deposited response labels. The data cannot establish whether a patient with a particular expression value would fare better with an alternative therapy.

2.2 Candidate selection and assay specificity

The panel was restricted to seven genes before examining their outcome associations. *BRCA1*, *BRCA2* and *RAD51* were selected for their roles in homologous recombination. *ERCC1* relates to nucleotide excision repair, *CCNE1* to cell-cycle regulation and replication stress, *SLFN11* to damage-associated replication arrest, and *ABCB1* to drug efflux. Table 1 records the evidence supporting these choices and the distinction between the measurement evaluated here and the assay that gives the biological interpretation.

Table 1. Candidate pathways and assay boundaries

Candidate	Biological rationale	Measurement limit of the present data
BRCA1 and *BRCA2*	Homologous recombination and restoration after secondary alterations	Transcript abundance does not identify pathogenic or restoring variants
RAD51	Repair activity, protein expression and nuclear foci	Bulk mRNA does not measure foci formation
CCNE1	Cell-cycle control and replication stress	Expression does not establish amplification
ERCC1	Nucleotide excision repair	Probe summaries do not resolve all isoforms or repair function
SLFN11	DNA damage-associated replication arrest	Bulk RNA is distinct from protein staining
ABCB1	Drug efflux and multidrug resistance	Combination treatment limits platinum-specific interpretation

Clinical studies of RAD51 have used quantitative protein imaging or nuclear foci rather than relying only on mRNA. Hoppe and colleagues studied quantitative RAD51 protein expression, while Compadre and colleagues evaluated RAD51 and DNA-damage foci in patient-derived models and tumour specimens [5, 10]. Functional homologous recombination studies have also examined primary cells and formalin-fixed material [23, 28]. A recent comparison of genomic and functional measurements further illustrates why an HRD result, RAD51 result and SLFN11 result should each retain its own assay definition [7]. The present data contain transcript measurements, so a weak transcript association does not directly test the performance of those assays.

SLFN11 has been evaluated by immunohistochemistry in ovarian cancer cohorts, including discovery and validation groups [1]. *CCNE1* amplification is a genomic event with recognised relevance to resistance biology, but bulk *CCNE1* expression can be influenced by cell-cycle composition and does not identify amplification by itself [14, 21]. *ABCB1* is relevant to multidrug resistance and was included as a broad resistance candidate. Its association in a combination-treated cohort cannot identify a platinum-specific transport effect. In particular, a taxane component may contribute to an observed association. Reviews of expression-based platinum-response signatures and PARP-inhibitor biomarkers show a wide range of candidate assays and validation strategies [27, 30]. Evidence about PARP-inhibitor benefit provides context for DNA repair biology, while the treatment endpoint must remain explicit.

3 Methods

3.1 Study design and public data sources

This retrospective secondary analysis used the deposited series matrices for GSE51373 and GSE63885 and the GPL570 annotation file downloaded on 19 November 2026. Both series use the Affymetrix Human Genome U133 Plus 2.0 array. The source studies were published by Koti and colleagues and Lisowska and colleagues [12, 15]. Their GEO records identify the sample sets, platform and clinical annotations [18, 19]. GEO provides a public archive for functional genomics data [2]. Only public de-identified records were used, and no contact with participants or access to clinical systems occurred.

GSE51373 contains 28 high-grade serous ovarian tumours. The source design compares twelve patients with progression-free survival below eight months with sixteen patients above eighteen months. The deposited labels are chemotherapy resistant and chemotherapy sensitive. All 28 labelled samples were retained. This is an extreme-outcome design and excludes the intervening range by source selection. The resistant category therefore differs from a uniform clinical six-month platinum-free interval rule.

GSE63885 contains 101 surgical tumour samples. The deposited histology field identifies 73 as serous, with the remainder endometrioid, clear cell or undifferentiated. Seventy-five records have an explicit sensitivity label. The primary external cohort required both a serous label and a sensitivity label, retaining 70 cases, of whom 32 were resistant. Three serous cases lacked a usable response label. The sensitivity analysis retained all 75 labelled tumours across histologies, with 34 resistant cases. The external serous cohort does not have a centrally reassessed modern high-grade classification in this analysis. It is therefore described as serous rather than treated as identical to the discovery population.

Table 2. Cohort selection and endpoint differences

Feature	Discovery GSE51373	External GSE63885
Deposited samples	28	101
Primary analysis	28 high-grade serous	70 labelled serous
Resistant and sensitive counts	12 and 16	32 and 38
Outcome source	PFS below 8 versus above 18 months	Deposited resistant versus moderate/high sensitivity
Excluded from primary analysis	None among deposited samples	28 non-serous, 3 unlabelled serous
Mixed-histology sensitivity	Not used	75 labelled, 34 resistant and 41 sensitive
Deposited preprocessing	MAS5, transformed log2(x + 1)	RMA, already logarithmic

The validation record describes resistant as disease-free survival below 180 days, moderately sensitive and highly sensitive as two longer-interval categories. The deposited text contains an inconsistent inequality for the intermediate category ("180>DFS>732"). This study retained the

explicit category assigned by the submitters and did not repair that string by inferring a new time boundary. Resistant was coded as one, and both sensitive categories as zero. Missing labels remained excluded. Overall survival, disease-free survival time and clinical response fields were not substituted for missing categories. The primary analysis thus evaluates the portability of the recorded labels, with the endpoint discrepancy retained as a limitation.

3.2 Processing and probe mapping

GSE51373 contains deposited MAS5 summaries on a nonnegative intensity scale, which were transformed as log2(intensity + 1). GSE63885 contains RMA-normalised data on an already logarithmic scale and was retained on that scale. The deposited processing descriptions and matrix distributions were checked before applying those rules. RMA is a standard approach for Affymetrix probe-level summarisation [11]. The current analysis uses the submitted series matrices and does not reprocess CEL files or perform joint normalisation across cohorts.

Platform rows were matched when the Gene symbol field exactly equalled the selected gene symbol. Rows mapping to multiple symbols were not included by this rule. For each gene and sample, the median of eligible probe-set values formed the expression summary. All genes were represented. *ABCB1* (209993_at, 243951_at), *BRCA1* (204531_s_at, 211851_x_at), *BRCA2* (208368_s_at, 214727_at), *CCNE1* (213523_at, 242105_at) and *RAD51* (205023_at, 205024_s_at) each had two probe sets, *ERCC1* had three (203719_at, 203720_s_at, 228131_at) and *SLFN11* had one (226743_at). No probe was selected by its association with response. The resulting seven-gene matrices contained no missing values.

Taking a median limits dependence on choosing a favourable probe but can combine probes that differ in specificity or transcript coverage. The preprocessing difference also means that an absolute numerical threshold in one cohort is not automatically transportable to the other. Rank-based discrimination is the primary outcome for that reason. The difference in group medians is shown within each cohort to document direction, with no interpretation as a directly comparable cross-cohort fold change.

3.3 Scoring and statistical analysis

For each gene, the raw discovery area under the receiver operating characteristic curve was calculated with resistance as the positive outcome. If the raw area was at least 0.5, expression itself was the score; otherwise the score was negative expression. To avoid floating-point ambiguity, this comparison was made on the integer Mann–Whitney count ($2U \geq n_1 n_0$), so that an exact area of 0.5 is a defined tie assigned to expression itself. The discovery-derived orientation was then fixed for external assessment. No external outcome was used to reverse or refit a score. *ERCC1* had an exact discovery tie (96 of 192 resistant–sensitive pairs), so it has no data-derived direction; its external area is reported for completeness but is not interpreted.

Discovery scores are optimistically oriented because the same labels determine direction and estimate performance. Their intervals are descriptive and conditional on the chosen orientation. The main inferential emphasis is external discrimination using the fixed orientation. The receiver operating characteristic area measures the probability that a randomly selected resistant case has a

higher score than a sensitive case, allowing for ties [9]. It does not quantify calibrated resistance risk or provide a useful clinical cut-off by itself. Precision and predictive values would additionally depend on prevalence and the selected operating point [24].

Percentile intervals used 1,500 stratified bootstrap resamples within each cohort, sampling resistant and sensitive cases with replacement while retaining their observed counts. A single random stream with seed 20261119 was used, processing genes in the order listed in Table 3 and, within each gene, the discovery cohort before the external cohort; resampling followed standard bootstrap principles [6]. The entire external score orientation remained fixed during resampling. Group differences were also tested with two-sided asymptotic Mann–Whitney tests. Benjamini–Hochberg correction was applied separately across the seven genes in each cohort, using a false discovery rate threshold of 0.05 [3]. The bootstrap intervals are pointwise and have no simultaneous coverage guarantee across genes. With 32 resistant and 38 sensitive external cases, the external intervals have a half-width of approximately 0.13, so the study can exclude strong discrimination but cannot distinguish modest association from none.

The analysis used Python with NumPy, pandas, SciPy and Matplotlib. Reporting was informed by principles in REMARK, STROBE and TRIPOD, with attention to assay definition, selection and external assessment [4, 16, 29]. No formal claim of checklist certification is made. PROBAST concepts informed the discussion of population and outcome differences [31]. The design includes no covariate adjustment, multivariable modelling, clinical utility analysis or treatment interaction test.

4 Results

4.1 Cohort accounting and observed expression signals

The primary analysis included 98 tumours across two cohorts. The discovery group contained 12 resistant and 16 sensitive samples. The external serous group contained 32 resistant and 38 sensitive samples. The mixed-histology sensitivity cohort added five labelled non-serous samples. Cohort composition and endpoint definitions are summarised in Table 2.

Discovery oriented areas ranged from 0.500 for *ERCC1* to 0.667 for *ABCB1*. *BRCA2* and *RAD51* each had an area of 0.656, *CCNE1* 0.651 and *SLFN11* 0.594. *BRCA1* was close to chance at 0.505. All discovery bootstrap intervals included 0.5 (Table 3). None of the seven discovery comparisons met the false discovery rate threshold; the smallest discovery q value was 0.325. The orientation favoured higher *ABCB1* and *BRCA1* and lower *BRCA2*, *CCNE1*, *RAD51* and *SLFN11*; *ERCC1* had no direction. These selected directions describe the cohort's rankings and cannot establish a coherent resistance mechanism.

4.2 External discrimination and direction changes

In the external serous cohort, oriented areas ranged from 0.434 for *CCNE1* to 0.615 for *ABCB1*. *ABCB1* had a 95% interval of 0.485–0.744 and an unadjusted group-comparison P value of approximately 0.100. Its adjusted value was approximately 0.700. The other adjusted values were

approximately 0.883. All seven pointwise external intervals included 0.5. Table 3 and Figure 1 report the fixed-direction results.

Table 3. Fixed-direction transcript discrimination

Gene	Discovery AUC	Discovery 95% CI	External AUC	External 95% CI	External q	Mixed-histology AUC
ABCB1	0.667	0.453–0.854	0.615	0.485–0.744	0.700	0.622
BRCA1	0.505	0.276–0.724	0.511	0.372–0.649	0.883	0.485
BRCA2	0.656	0.427–0.865	0.469	0.338–0.607	0.883	0.489
CCNE1	0.651	0.422–0.849	0.434	0.292–0.580	0.883	0.460
ERCC1[a]	0.500	0.266–0.708	0.441	0.304–0.575	0.883	0.456
RAD51	0.656	0.422–0.870	0.480	0.348–0.611	0.883	0.488
SLFN11	0.594	0.370–0.802	0.473	0.329–0.604	0.883	0.477

Discovery, GSE51373 (n = 28); external, GSE63885 serous (n = 70); mixed histology, all labelled GSE63885 tumours (n = 75). Intervals are pointwise percentile intervals from 1,500 stratified bootstrap resamples; q values are Benjamini–Hochberg adjusted.
[a] Exact discovery tie; oriented by the tie rule (expression itself). Under the reverse orientation the external AUC would be 0.559. Not interpreted.

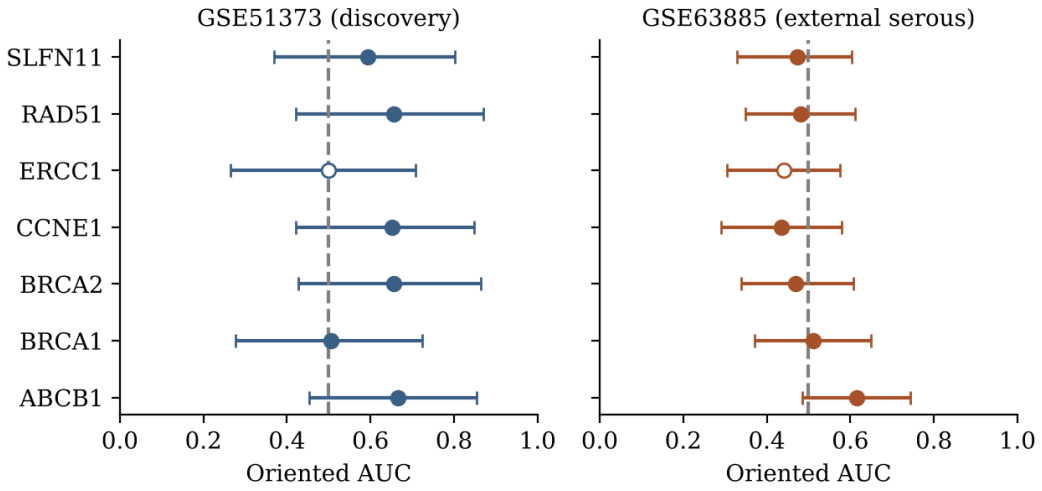

Figure 1. Discovery and external AUCs for the seven transcript summaries. External directions are fixed from discovery. Whiskers are pointwise percentile 95% intervals from 1,500 stratified bootstrap resamples. Discovery estimates are optimistically oriented. The open marker for *ERCC1* indicates an exact discovery tie, oriented by the prespecified tie rule.

BRCA2, *CCNE1*, *RAD51* and *SLFN11* had external areas below 0.5 under their discovery-selected score orientation, and *BRCA1* remained close to chance. *ERCC1* is excluded from this comparison because it had no discovery direction. Figure 2 shows the group median differences underlying part of this inconsistency. For example, the discovery median difference for *BRCA2* was negative, while the external difference was positive. *SLFN11* had a near-zero positive median difference in

discovery despite a negative score orientation selected from its full rank distribution. The median alone and the area under the curve therefore describe different aspects of the sample distributions.

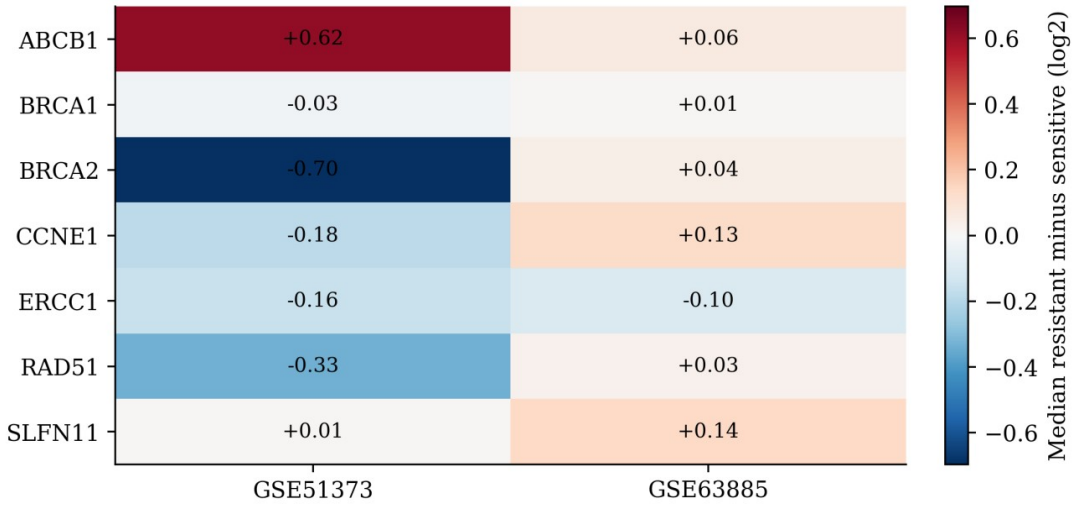

Figure 2. Within-cohort median expression in resistant minus sensitive groups. Values are on their respective processed log2 scales. Differences in source preprocessing prevent direct interpretation of cross-cohort magnitudes as matched fold changes.

SLFN11 was selected for a separate receiver operating characteristic plot because of its recent ovarian-cancer biomarker literature. It was not chosen as the highest-performing transcript after analysis. Figure 3 shows discovery and external curves using the same discovery-selected direction. The external curve illustrates weak discrimination across available thresholds. Neither its shape nor the discovery curve supports choosing a clinical decision threshold from this sample.

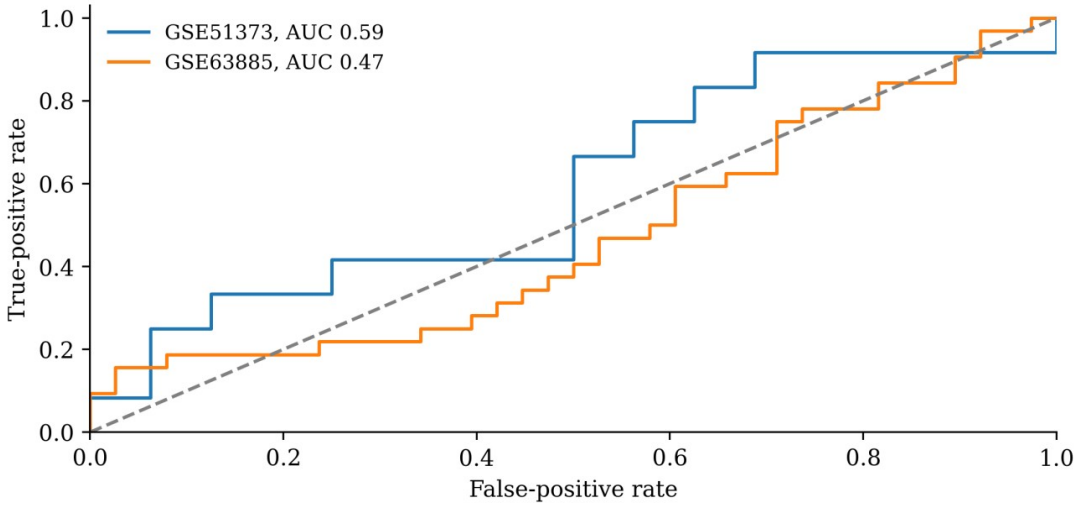

Figure 3. *SLFN11* receiver operating characteristic curves with the discovery-selected score direction. *SLFN11* was selected for display on biological grounds. No operating threshold was selected.

4.3 Histology sensitivity

Including all 75 labelled GSE63885 tumours produced an *ABCB1* area of approximately 0.622. Other oriented areas remained close to 0.5 or below it (Table 3). Figure 4 compares these estimates

with the primary serous-only results. Adding the five labelled non-serous tumours did not yield a clearly useful discriminator in this panel. This modest change cannot establish histology independence because the additional group is small and the histological categories are heterogeneous.

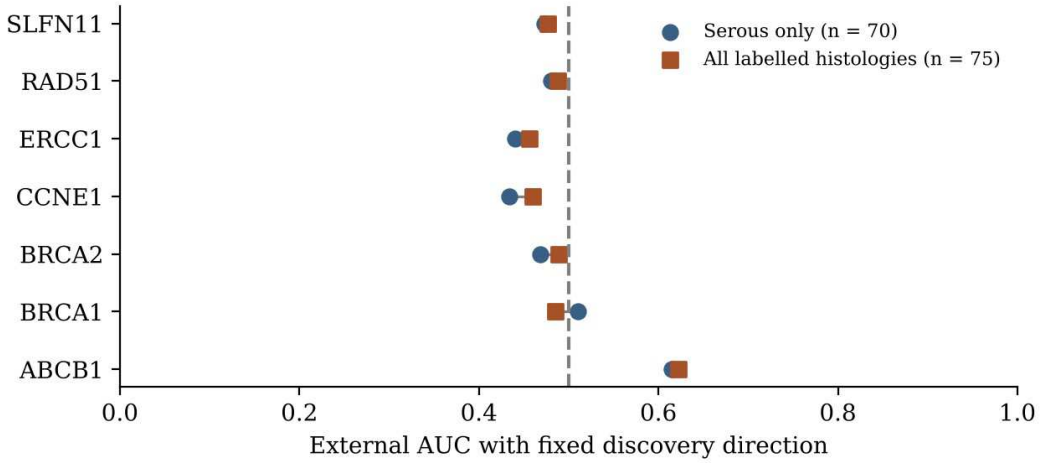

Figure 4. External AUC sensitivity to histology inclusion. Serous-only estimates use 70 tumours. Mixed-histology estimates use all 75 labelled tumours. All directions remain fixed from discovery.

5 Discussion

5.1 What the reanalysis establishes

The evaluated transcripts gave little evidence of portable discrimination between the deposited response categories. The result is useful for understanding the gap between pathway plausibility and assay performance. DNA repair and damage tolerance remain relevant biological mechanisms, but the present gene-level summaries did not consistently order resistant and sensitive cases across the two cohorts. The external uncertainty intervals are also wide enough to accommodate some modest association. Accordingly, the study identifies limited evidence under this design and does not establish the absence of a biologically meaningful role.

ABCB1 retained the strongest numerical area in the external analysis, yet its interval crosses 0.5 and the multiple-testing result provides little support for a reproducible association. It should not be promoted to a validated marker on the basis of its rank among seven weak candidates. In a combination-treated cohort, any association could also relate to the co-administered agent or to a broader resistant-cell state. A mechanistic claim about platinum transport would require a different experimental design and direct pharmacological evidence.

The results for *BRCA1* and *BRCA2* are equally important to interpret at the correct assay level. Many clinically relevant alterations are pathogenic sequence changes, copy alterations, promoter methylation states or restoring mutations. Median probe expression cannot resolve those events. The same issue applies to RAD51 foci and functional homologous recombination assays. The present results are compatible with those assays carrying useful information while a bulk mRNA

summary performs weakly. Table 4 states the next measurements needed to address the biological and clinical questions.

Table 4. Validation questions following the transcript reanalysis

Question	Required evidence	Interpretive purpose
Which molecular state is measured?	Validated variant, amplification, methylation, protein or functional assay	Match the readout to the proposed resistance mechanism
Which clinical outcome is predicted?	Specified treatment line, time origin and adjudicated outcome	Reduce endpoint mismatch
Does the signal add information?	Locked analysis with clinical covariates and external testing	Estimate incremental value
Can the result guide treatment?	Treatment comparison, calibration and clinical impact evidence	Assess differential benefit and clinical utility

5.2 Endpoint transport and study design

GSE51373 deliberately separates extreme progression-free survival groups. GSE63885 uses broader deposited sensitivity categories with a different time reference. A gene can perform differently when the endpoint includes different clinical processes or includes patients near a boundary that the discovery design omitted. The original GSE63885 publication itself emphasised histological variation and problems with reproducibility [15]. These concerns remain pertinent when public datasets are reused for new biomarker signatures. Explicitly checking labels can prevent an apparently simple binary analysis from concealing a change in what counts as resistance.

The absence of treatment comparison is a further limit. A strong response association would still require investigation of whether the biomarker modifies the benefit of platinum relative to another approach. Prognosis and treatment sensitivity can share determinants, especially when recurrence timing is used to define outcome. A prospective study could specify the intended clinical decision, sampling time, treatment line, outcome window and assay before enrolment. It could then evaluate whether the marker improves decisions beyond treatment history, histology, residual disease and other available information.

5.3 Laboratory translation and validation priorities

For laboratory use, the assay should match the mechanism being proposed. A BRCA analysis requires validated variant interpretation and an appropriate specimen. Functional RAD51 assessment requires evidence of an interpretable proliferating, DNA-damage-positive tumour-cell population and an established scoring procedure. SLFN11 immunohistochemistry requires a reproducible staining and scoring protocol with clear handling of intratumour heterogeneity. Gene-expression measurements require documented RNA quality, tumour content, transcript specificity and handling of platform effects.

Archived-specimen validation requires an explicit intended use, reliable specimen handling and a design that protects against selective analysis [25]. A useful next step would collect those readouts

in the same cohort, with blinded laboratory scoring and an independently adjudicated platinum outcome. The comparison should test whether an expression summary adds information to the direct molecular or functional measurement. Confidence intervals, missingness and assay failure should be reported alongside discrimination. If risk prediction is intended, a locked model would also need calibration and a justified decision threshold in an external population. A change in treatment based on that score would require evidence of clinical benefit.

6 Limitations

The discovery sample is small and selected by extreme outcomes. External patients differ in recorded histology, endpoint construction and treatment context. The public series matrices have different preprocessing, and this study does not reconstruct both cohorts from raw CEL files through a common pipeline. Although within-cohort rankings limit dependence on numerical scale, the median across multiple probes can still be affected by summarisation and probe specificity. The panel was motivated by literature, but there was no prospectively registered protocol or blinded external analyst.

Bulk tumour expression also mixes cell populations. A transcript may reflect immune or stromal cells, tumour purity, grade or proliferation. No adjustment was made for age, residual disease, stage, regimen or BRCA mutation status, although GSE63885 deposits residual disease, *BRCA1* mutation status and the adjuvant regimen (platinum with cyclophosphamide or with a taxane). Those fields were not used, so the analysis cannot establish independent association or incremental value, and the taxane question raised for *ABCB1* remains untested. Bootstrap intervals quantify sampling variation within each recorded dataset and exclude uncertainty introduced by cohort selection, annotation and the difference between historical labels and present clinical definitions.

The data do not include the direct assays needed to test HRD, functional repair, *CCNE1* amplification or SLFN11 protein expression as clinical predictors. Nor does the study evaluate longitudinal changes after treatment, including BRCA reversions. Findings should be limited to the seven transcript summaries and the selected deposited outcomes. They provide no basis to withhold platinum, select another treatment or change a patient's clinical management.

7 Conclusion

Seven biologically motivated transcripts showed weak and inconsistent external discrimination of platinum-response labels in this public-cohort reanalysis. No gene met the false discovery rate threshold, and the highest external area had substantial uncertainty. The findings support keeping molecular mechanism, assay readout and clinical endpoint aligned when developing biomarkers of platinum resistance. Further validation should use harmonised outcomes and direct molecular or functional measurements relevant to the proposed mechanism.

Declarations

Data and code availability. All data are publicly available from the NCBI Gene Expression Omnibus under accessions GSE51373 and GSE63885, with platform annotation GPL570. The reproduction script, which reads the deposited series matrices and annotation file and regenerates

every value in Table 3 and Figures 1 and 4, is available at [GitHub repository URL] and archived at [Zenodo DOI].

Ethics statement. This study used only publicly deposited, de-identified data. No human participants were recruited and no identifiable records were accessed; ethical approval and informed consent were therefore not required.

Funding. [Confirm: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.]

Competing interests. [Confirm: The author declares no competing interests.]

Author contributions. F.A. conceived the study, performed the analysis, prepared the figures and wrote and revised the manuscript.

Use of AI tools. [State here whether and how AI tools were used, if the target journal requires it.]

References

[1] Akashi, H., Yachida, N., Ueda, H., Yamaguchi, M., Yamawaki, K., Tamura, R., et al. (2024). SLFN11 is a BRCA independent biomarker for the response to platinum-based chemotherapy in high-grade serous ovarian cancer and clear cell ovarian carcinoma. *Molecular Cancer Therapeutics*, 23(1), 106–116. https://doi.org/10.1158/1535-7163.MCT-23-0257

[2] Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Research*, 41(Database issue), D991–D995. https://doi.org/10.1093/nar/gks1193

[3] Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, 57(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

[4] Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Annals of Internal Medicine*, 162(1), 55–63. https://doi.org/10.7326/M14-0697

[5] Compadre, A. J., van Biljon, L. N., Valentine, M. C., Llop-Guevara, A., Graham, E., Fashemi, B., et al. (2023). RAD51 foci as a biomarker predictive of platinum chemotherapy response in ovarian cancer. *Clinical Cancer Research*, 29(13), 2466–2479. https://doi.org/10.1158/1078-0432.CCR-22-3335

[6] Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman & Hall. https://doi.org/10.1201/9780429246593

[7] Feng, Z., Zhu, C., Zhang, X., Huang, Z., Ju, X., Guo, Q., et al. (2025). Comprehensive evaluation of genomic and functional assays for homologous recombination deficiency with high-grade epithelial ovarian cancer: platinum sensitivity and prognosis. *International Journal of Gynecological Cancer*, 35(1), 100031. https://doi.org/10.1016/j.ijgc.2024.100031

[8] González-Martín, A., Ledermann, J. A., & ESMO Guidelines Committee. (2026). ESMO Clinical Practice Guideline Express Update on the management of epithelial ovarian cancer. *ESMO Open*, 11(2), 106032. https://doi.org/10.1016/j.esmoop.2025.106032

[9] Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), 29–36. https://doi.org/10.1148/radiology.143.1.7063747

[10] Hoppe, M. M., Jaynes, P., Wardyn, J. D., Upadhyayula, S. S., Tan, T. Z., Lie, S., et al. (2021). Quantitative imaging of RAD51 expression as a marker of platinum resistance in ovarian cancer. *EMBO Molecular Medicine*, 13(5), e13366. https://doi.org/10.15252/emmm.202013366

[11] Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B., & Speed, T. P. (2003). Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Research*, 31(4), e15. https://doi.org/10.1093/nar/gng015

[12] Koti, M., Gooding, R. J., Nuin, P., Haslehurst, A., Crane, C., Weberpals, J., et al. (2013). Identification of the IGF1/PI3K/NF κB/ERK gene signalling networks associated with chemotherapy resistance and treatment response in high-grade serous epithelial ovarian cancer. *BMC Cancer*, 13, 549. https://doi.org/10.1186/1471-2407-13-549

[13] Lheureux, S., Braunstein, M., & Oza, A. M. (2019). Epithelial ovarian cancer: evolution of management in the era of precision medicine. *CA: A Cancer Journal for Clinicians*, 69(4), 280–304. https://doi.org/10.3322/caac.21559

[14] Lheureux, S., Gourley, C., Vergote, I., & Oza, A. M. (2019). Epithelial ovarian cancer. *Lancet*, 393(10177), 1240–1253. https://doi.org/10.1016/S0140-6736(18)32552-2

[15] Lisowska, K. M., Olbryt, M., Dudaladava, V., Pamuła-Piłat, J., Kujawa, K., Grzybowska, E., et al. (2014). Gene expression analysis in ovarian cancer—faults and hints from DNA microarray study. *Frontiers in Oncology*, 4, 6. https://doi.org/10.3389/fonc.2014.00006

[16] McShane, L. M., Altman, D. G., Sauerbrei, W., Taube, S. E., Gion, M., Clark, G. M., et al. (2005). Reporting recommendations for tumor marker prognostic studies (REMARK). *Journal of the National Cancer Institute*, 97(16), 1180–1184. https://doi.org/10.1093/jnci/dji237

[17] Miller, R. E., Leary, A., Scott, C. L., Serra, V., Lord, C. J., Bowtell, D., et al. (2020). ESMO recommendations on predictive biomarker testing for homologous recombination deficiency and PARP inhibitor benefit in ovarian cancer. *Annals of Oncology*, 31(12), 1606–1622. https://doi.org/10.1016/j.annonc.2020.08.2102

[18] National Center for Biotechnology Information. GSE51373: Gene expression data from high grade serous ovarian cancer. Gene Expression Omnibus. Retrieved 19 November 2026, from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE51373

[19] National Center for Biotechnology Information. GSE63885: Gene expression profiling in ovarian cancer. Gene Expression Omnibus. Retrieved 19 November 2026, from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE63885

[20] Norquist, B., Wurz, K. A., Pennil, C. C., Garcia, R., Gross, J., Sakai, W., et al. (2011). Secondary somatic mutations restoring BRCA1/2 predict chemotherapy resistance in hereditary ovarian carcinomas. *Journal of Clinical Oncology*, 29(22), 3008–3015. https://doi.org/10.1200/JCO.2010.34.2980

[21] Patch, A. M., Christie, E. L., Etemadmoghadam, D., Garsed, D. W., George, J., Fereday, S., et al. (2015). Whole-genome characterization of chemoresistant ovarian cancer. *Nature*, 521(7553), 489–494. https://doi.org/10.1038/nature14410

[22] Pennington, K. P., Walsh, T., Harrell, M. I., Lee, M. K., Pennil, C. C., Rendi, M. H., et al. (2014). Germline and somatic mutations in homologous recombination genes predict platinum response and survival in ovarian, fallopian tube, and peritoneal carcinomas. *Clinical Cancer Research*, 20(3), 764–775. https://doi.org/10.1158/1078-0432.CCR-13-2287

[23] Pikkusaari, S., Tumiati, M., Virtanen, A., Oikkonen, J., Li, Y., Perez-Villatoro, F., et al. (2023). Functional homologous recombination assay on FFPE specimens of advanced high-grade serous ovarian cancer predicts clinical outcomes. *Clinical Cancer Research*, 29(16), 3110–3123. https://doi.org/10.1158/1078-0432.CCR-22-3156

[24] Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3), e0118432. https://doi.org/10.1371/journal.pone.0118432

[25] Simon, R. M., Paik, S., & Hayes, D. F. (2009). Use of archived specimens in evaluation of prognostic and predictive biomarkers. *Journal of the National Cancer Institute*, 101(21), 1446–1452. https://doi.org/10.1093/jnci/djp335

[26] Swisher, E. M., Sakai, W., Karlan, B. Y., Wurz, K., Urban, N., & Taniguchi, T. (2008). Secondary BRCA1 mutations in BRCA1-mutated ovarian carcinomas with platinum resistance. *Cancer Research*, 68(8), 2581–2586. https://doi.org/10.1158/0008-5472.CAN-08-0088

[27] Talijanovic, M., Lopacinska-Jørgensen, J., & Høgdall, E. V. (2025). Gene expression profiles in ovarian cancer tissues as a potential tool to predict platinum-based chemotherapy resistance. *Anticancer Research*, 45(1), 11–26. https://doi.org/10.21873/anticanres.17389

[28] Tumiati, M., Hietanen, S., Hynninen, J., Pietilä, E., Färkkilä, A., Kaipio, K., et al. (2018). A functional homologous recombination assay predicts primary chemotherapy response and long-term survival in ovarian cancer patients. *Clinical Cancer Research*, 24(18), 4482–4493. https://doi.org/10.1158/1078-0432.CCR-17-3770

[29] von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., Vandenbroucke, J. P., et al. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Annals of Internal Medicine*, 147(8), 573–577. https://doi.org/10.7326/0003-4819-147-8-200710160-00010

[30] Wang, Y. W., Allen, I., Funingana, G., Tischkowitz, M., & Joko-Fru, Y. W. (2025). Predictive biomarkers for the efficacy of PARP inhibitors in ovarian cancer: an updated systematic review. *BJC Reports*, 3(1), 14. https://doi.org/10.1038/s44276-025-00122-9

[31] Wolff, R. F., Moons, K. G. M., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., et al. (2019). PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Annals of Internal Medicine*, 170(1), 51–58. https://doi.org/10.7326/M18-1376