Yang *et al. Diagnostic Pathology* (2025) 20:117
https://doi.org/10.1186/s13000-025-01726-y

Diagnostic Pathology

RESEARCH Open Access

Clinical significance and expression of ALDH1 in triple-negative breast cancer

Yaoyao Yang[1,2,3,4,5], Zhuoran Li[1,2,3,4,5], Yangyang Zhu[1,2,3,4,5], Chongyong Guo[6], Li Yang[7] and Jian Zhang[1,2,3,4,5*]

Abstract

Background Triple-negative breast cancer (TNBC) is aggressive and has limited therapeutic options due to the absence of targeted therapies, highlighting the urgent need for prognostic biomarkers linked to cancer stemness and chemoresistance. Aldehyde dehydrogenase 1 (ALDH1), a key regulator of stem cell properties, remains incompletely characterized in TNBC clinical cohorts.

Methods ALDH1 mRNA expression levels were analyzed using the GEO2R online database, and its prognostic significance was assessed via the Kaplan–Meier plotter tool. Immunohistochemical (IHC) staining was performed on a tissue microarray comprising 96 TNBC samples and paired adjacent normal tissues from patients treated at Binzhou People's Hospital between 2016 and 2022. The associations between ALDH1 expression and clinicopathological parameters were evaluated using the chi-square test.

Results Bioinformatics analysis revealed significantly higher ALDH1 mRNA expression in TNBC tissues compared to adjacent benign tissues. Kaplan–Meier survival analysis demonstrated that elevated ALDH1 mRNA expression was associated with poor prognosis in TNBC patients. IHC staining further confirmed elevated ALDH1 protein expression in TNBC tissues compared with normal adjacent tissues. However, there was no significant correlation between ALDH1 expression and conventional clinicopathological parameters, including age, menopausal status, tumor size, TNM stage, histological grade, histological subtype, axillary lymph node metastasis and the Ki-67 index ($p > 0.05$). High ALDH1 expression was significantly associated with poorer overall survival ($\chi^2 = 16.836$, $p < 0.001$).

Conclusion Our data demonstrate that ALDH1 expression is not significantly associated with conventional clinicopathological parameters (such as age, TNM stage, or histological grade). Instead, it is associated with poorer survival on univariate analysis in TNBC patients. Its lack of association with clinicopathological factors suggests its potential utility as a supplementary prognostic indicator.

Keywords TNBC, ALDH1, IHC, TMA

*Correspondence:
Jian Zhang
2269776489@qq.com
[1]Department of Central Laboratory, Binzhou People's Hospital, Binzhou 256600, People's Republic of China
[2]Shandong Provincial Key Laboratory of Diagnosis and Treatment for Infectious Diseases and Medicine, Health of Binzhou People's Hospital, Binzhou 256600, People's Republic of China
[3]Binzhou Key Laboratory of Pathogenic Biology of Binzhou People's Hospital, Binzhou 256600, People's Republic of China
[4]Binzhou Engineering Research Center of Precision Medicine and Regenerative Medicine, Binzhou People's Hospital, Binzhou 256600, People's Republic of China
[5]Binzhou Engineering Research Center of Zoonotic Diseases, Binzhou People's Hospital, Binzhou 256600, People's Republic of China
[6]Department of Breast Center, Binzhou People's Hospital, Binzhou 256600, People's Republic of China
[7]Department of Laboratory Medicine, Binzhou People's Hospital, Binzhou 256600, People's Republic of China

Introduction

Breast cancer is the most common cancer in women worldwide. According to 2022 global estimates, approximately 2.3 million new cases of breast cancer were diagnosed worldwide, resulting in 670,000 breast cancer-related deaths [1, 2]. Triple-negative breast cancer (TNBC), defined by the absence of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) expression [1], accounts for 15–20% of all breast cancer cases and demonstrates distinct molecular heterogeneity and aggressive clinical behavior [3, 4]. TNBC patients face a 40% higher risk of distant recurrence within the first 3 years compared to hormone receptor-positive subtypes [4, 5]. The lack of targeted therapies leaves platinum-based chemotherapy as the primary treatment, yet 30–50% of patients develop resistance, leading to inevitable progression [6, 7]. This unmet clinical need underscores the urgency of identifying novel therapeutic vulnerabilities rooted in TNBC biology.

A key contributor to TNBC aggressiveness and therapeutic resistance is the presence of cancer stem cells (CSCs), a subpopulation of tumor cells with self-renewal capacity and differentiation potential. CSCs are implicated in tumor initiation, metastasis, and relapse due to their resistance to conventional therapies [8]. Breast cancer CSCs exhibit dynamic phenotypic plasticity, enabling bidirectional transitions among quiescent, invasive, mesenchymal, and highly proliferative epithelial-like states [9, 10]. In TNBC, unique molecular pathways and tumor microenvironment components collaboratively sustain the CSCs phenotype. TNBC-CSCs can maintain their survival via multiple mechanisms, including the synthesis of drug resistance-associated proteins, the activation of DNA damage repair pathways, the suppression of apoptotic signaling cascades, and the induction of protective autophagy [11]. CSCs exhibit self-renewal capacity through Wnt/β-catenin and Notch pathway activation [12]. Lu et al. reported that epithelial–mesenchymal transition (EMT) in CSCs upregulates CD90 expression, which subsequently activates the Src and nuclear factor-κB (NF-κB) signaling pathways in tumor cells [13]. This activation induces CSCs secretion of cytokines, including interleukin-6 (IL-6) and interleukin-8 (IL-8), which reinforce stemness maintenance and functional enhancement of CSCs [12].

Aldehyde dehydrogenase 1 (ALDH1) has emerged as a critical functional biomarker and metabolic regulator in TNBC-CSCs, serving as a nexus between stemness maintenance and therapeutic resistance. High ALDH1 activity is strongly associated with CSCs properties, including enhanced tumorigenicity and metastatic potential [14–18]. ALDH1 synergistically sustains self-renewal capacity through epigenetic modulation of pluripotency factors [19, 20]. ALDH1 catalyzes retinoic acid synthesis and detoxifies reactive aldehydes, thereby maintaining stemness via ROS reduction and DNA damage repair [20]. ALDH1 catalyzes the oxidation of intracellular aldehydes, thereby protecting cells from oxidative stress and conferring chemoresistance [21]. Marcato et al. demonstrated that elevated ALDH1A3 expression in breast cancer patient tumors and cell lines is correlated with poorer prognosis and TNBC subtypes, driving tumor progression through the activation of retinoic acid (RA) signaling pathways [22]. ALDH1 reinforces a stem-like phenotype in TNBC by forming a positive feedback loop with core stemness pathways and by directly driving the EMT [22–24]. Despite these findings, the molecular mechanisms by which ALDH1 regulates TNBC stemness and progression remain incompletely understood.

This study investigated the expression levels of ALDH1 in TNBC through RNA bioinformatics analysis and subsequent validation via immunohistochemical (IHC) analysis in a clinically annotated TNBC cohort ($n = 96$). We systematically evaluated the associations between ALDH1 expression and clinical outcomes to explore its potential as a prognostic biomarker and therapeutic target.

Materials and methods

Bioinformatics analysis

Transcriptome data from the GEO database (www.ncbi.nlm.nih.gov/geo, GSE38959 and GSE52194) were analyzed to compare ALDH1 mRNA expression levels between TNBC tissues and adjacent normal tissues. The prognostic value of ALDH1 mRNA expression in TNBC was subsequently assessed using the Kaplan–Meier plotter tool (www.kmplot.com). Patients were stratified into 'high' and 'low' expression groups based on the best-performing threshold automatically calculated by the tool's algorithm [25].

Patients and samples

This study utilized a well-characterized cohort ($n = 96$) of TNBC patients who underwent treatment at Binzhou People's Hospital between 2016 and 2022 (see Additional file 1). All participants were treatment-naive and had not received neoadjuvant radiotherapy, chemotherapy, or biological immunotherapy prior to surgical intervention. Clinicopathological parameters and prognostic data were retrospectively collected from patients' medical records. Tissue microarrays (TMA) were constructed containing archival, formalin-fixed and paraffin-embedded (FFPE) materials from surgically resected breast cancer specimens. Each TMA core (2 mm in diameter) contained paired samples of malignant lesions with immediately adjacent normal tissue. The tumor tissues were histologically diagnosed and classified using the World Health

Organization (WHO) classification of breast tumors. Written informed consent was obtained from all participants prior to their participation in the study.

Immunohistochemistry (IHC) of tissue microarray assay

Formalin-fixed and paraffin-embedded TMA sections were deparaffinized and rehydrated using xylene and a graded alcohol series. Antigen retrieval was performed via heat-induced epitope retrieval (HIER) at 97 °C for 30 min using EDTA buffer (pH 9.0) in a temperature-controlled water bath (DAKO, Denmark). Endogenous peroxidase was blocked with 0.3% hydrogen peroxide for 10 min. The sections were incubated with anti-ALDH1 primary antibody (1:200 dilution, BD Biosciences, Cat# 611195) for 1 h at room temperature. Antibody binding was visualized using a commercial DAB detection kit (ZSGB-BIO, Beijing, China; Cat# ZLI-9018). Digital images of all the stained sections were acquired using a high-resolution whole slide scanner (KFBIO, Ningbo, China).

Immunostostaing analysis

All slides were independently evaluated by two senior pathologists who were blind to the patients' clinical data, and the final staining score was calculated based on the intensity and percentage of positive cancer cells and normal glandular cells. The staining intensity was classified into 4 levels: 0 (no staining), 1 + (mild staining), 2 + (moderate staining), and 3 + (intense staining). For ALDH1, the percentage of positive cells and the H-score were obtained (0–300) [26]. The optimal cutoff value of the H-score was set as 95 using X-tile software (version: 3.6.1, Yale University) based on the patients' 5-year survival time [27]. Cases with scores from 0 to 95 were considered as low cytoplasmic expression, and scores from 96 to 300 considered positive/high cytoplasmic expression.

Statistical analysis

The associations between ALDH1 expression and clinicopathological characteristics were assessed using Pearson's chi-square test. Survival outcomes were analyzed via Kaplan–Meier curves with log-rank tests for group comparisons. All the statistical tests were two-sided, with a p value < 0.05 considered statistically significant. All analyses were performed via GraphPad Prism software (version 10.0, USA).

Results

Elevated ALDH1 mRNA expression and its prognostic significance in TNBC

Bioinformatic analysis of transcriptome data from the GEO database revealed significantly elevated ALDH1 mRNA expression levels in TNBC tissues compared with adjacent benign tissues (Fig. 1). Kaplan–Meier survival analysis further demonstrated that high ALDH1 mRNA expression was associated with poorer overall survival in TNBC patients (HR = 2.91, 95% CI: 1.21–6.96, $p = 0.012$) (Fig. 2).

IHC Validation of ALDH1 expression

Initial immunohistochemical analysis encompassed a cohort of 96 TNBC samples with paired adjacent normal tissues. Following rigorous histological quality control, 44 cases (45.8%) were excluded from paired comparative analysis due to: (i) procedural loss of tissue sections during microtomy processing ($n = 20$), predominantly

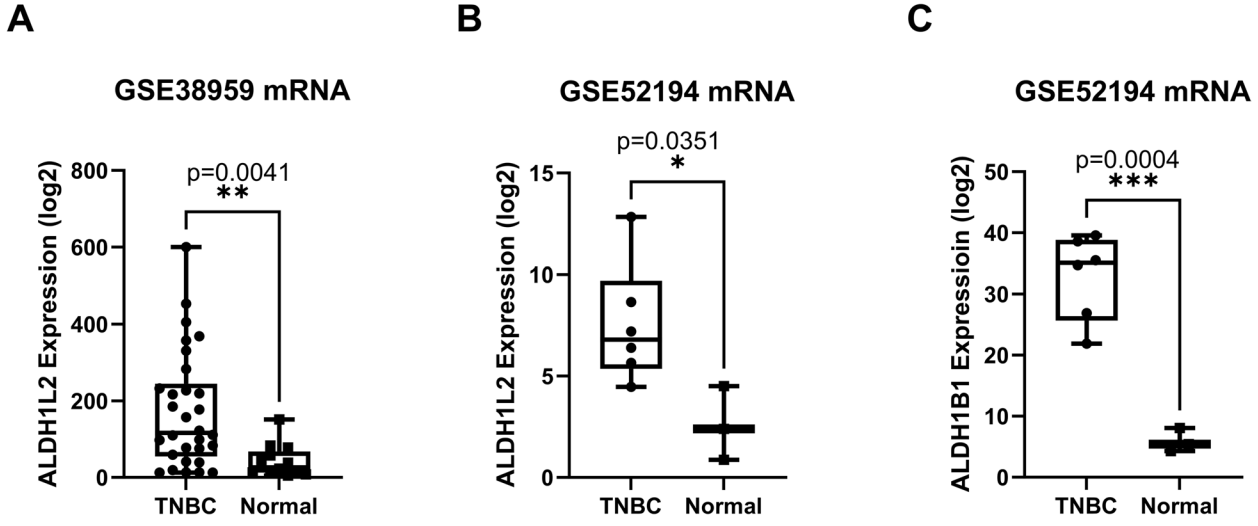

Fig. 1 Upregulation of ALDH1 mRNA in TNBC tissues revealed by GEO transcriptome analysis. **A** Analysis of the GSE38959 dataset revealed significantly higher ALDH1L2 mRNA expression in TNBC tissues than in adjacent normal tissues. **B** Analysis of the GSE52194 dataset revealed significantly higher ALDH1L2 mRNA expression in TNBC tissues than in adjacent normal tissues. **C** Analysis of the GSE52194 dataset revealed significantly higher ALDH1B1 mRNA expression in TNBC tissues than in adjacent normal tissues

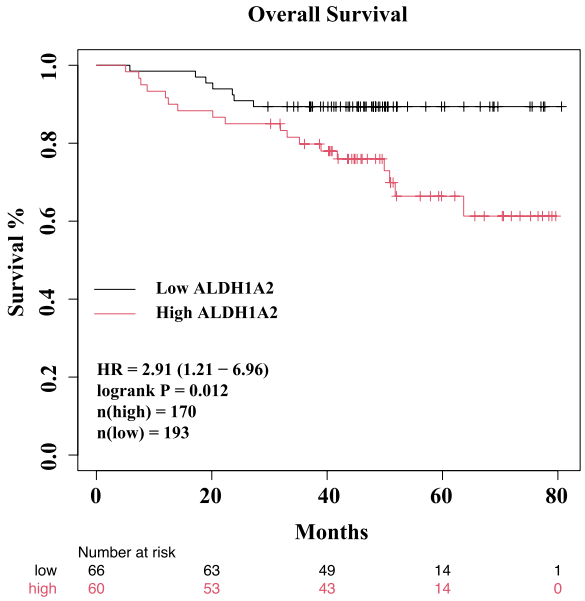

Fig. 2 Kaplan–Meier survival analysis of TNBC patients based on ALDH1 mRNA expression ($p = 0.012$). Kaplan–Meier analysis shows significantly reduced overall survival in TNBC patients with high ALDH1 mRNA expression compared to low expression counterparts (HR = 2.91, 95% CI: 1.21–6.96; log-rank $P = 0.012$)

affecting adipose-replaced stromal compartments in adjacent tissues, and (ii) insufficient glandular content in histologically defined "normal" adjacent tissues (< 10% lobular epithelium, $n = 24$). The final evaluable cohort comprised 52 matched TNBC-normal tissue pairs. Despite these isolated imbalances, the overall cohort remained well-balanced. However, we acknowledge these imbalances as a limitation of the paired analysis and considered them in the interpretation of the outcomes (see Additional file 2). Subsequent survival analysis and

Table 1 Comparison of ALDH1 expression between TNBC tissues and paired adjacent normal tissues ($N = 52$)

	Normal: ALDH1 high n (%)	Normal: ALDH1 low n (%)	Total N (%)
TNBC: ALDH1 high n (%)	7 (13.5%)	23 (44.2%)	30 (57.7%)
TNBC: ALDH1 low n (%)	2 (3.8%)	20 (38.5%)	22 (42.3%)
Total N (%)	9 (17.3%)	43 (82.7%)	52 (100%)

McNemar's χ^2 test: $\chi^2 = 17.64$, $p < 0.001$

clinicopathological parameter assessments were based on the complete cohort of 96 samples. IHC analysis of 52 TNBC samples revealed differential spatial expression of ALDH1. Among the TNBC samples, high ALDH1 immunoreactivity was detected in 30 cases (Fig. 3D), whereas low expression was observed in 22 cases (Fig. 3C). In contrast, adjacent normal tissues exhibited low ALDH1 expression in 43 cases (Fig. 3A), with only 9 cases demonstrating high immunoreactivity (Fig. 3B). Corresponding hematoxylin and eosin (H&E) staining confirmed the histological integrity of all representative tissues (Fig. 3E–H). A quantitative H-score assessment revealed a greater proportion of ALDH1-high samples in TNBC tissues (30/52, 57.7%) than in paired adjacent normal tissues (9/52, 17.3%) (McNemar's $\chi^2 = 17.64$, $p < 0.001$; Table 1).

Clinicopathological correlation analysis

Univariate assessment revealed no statistically significant associations between ALDH1 protein expression (high vs. low) and conventional clinicopathological parameters, including patient demographics (age, menopausal status), tumor characteristics (TNM stage, histological grade/

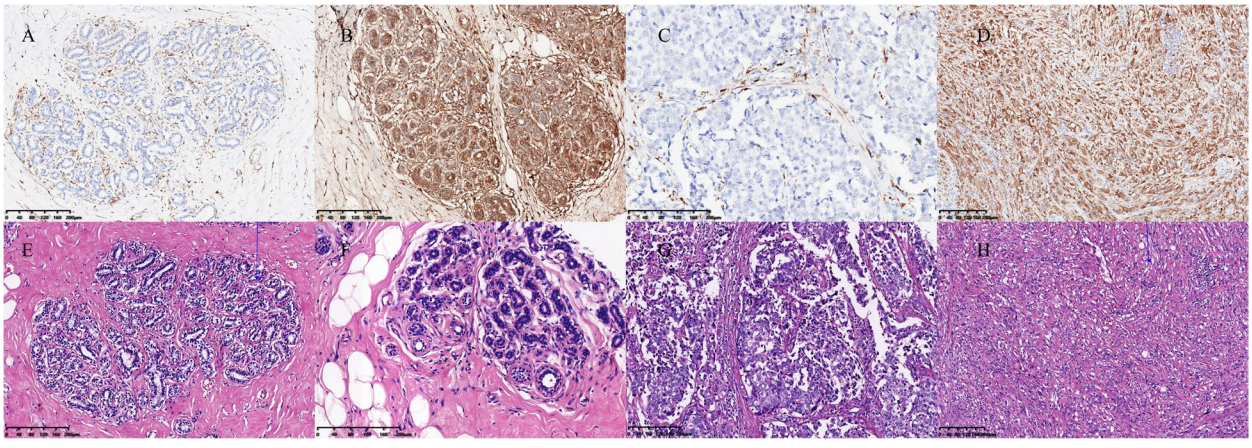

Fig. 3 ALDH1 expression in TNBC and paired adjacent normal tissues. **A** Adjacent normal tissue with low ALDH1 immunoreactivity. **B** Adjacent normal tissue with high ALDH1 immunoreactivity. **C** TNBC tissue with low ALDH1 immunoreactivity. **D** TNBC tissue with high ALDH1 immunoreactivity. **E** H&E staining of the adjacent normal tissue region shown in panel A, confirming normal histological architecture. **F** H&E staining of the adjacent normal tissue region shown in panel B, confirming normal histological architecture. **G** H&E staining of the TNBC region shown in panel C, confirming malignant histology. **H** H&E staining of the TNBC region shown in panel D, confirming malignant histology. Scale bars: 200 µm (applicable to all panels)

size/subtype), or proliferation indices (Ki-67). All comparisons yielded $p > 0.05$ (Table 2).

ALDH1 overexpression predicts adverse survival outcomes

ALDH1 expression was significantly associated with mortality in univariate analysis. Patients with high ALDH1 expression (H-score ≥ 95; $n = 31$) exhibited a higher mortality rate (15/31, 48.4%). In contrast, patients with low ALDH1 expression (H-score < 95; $n = 65$) showed significantly lower mortality (7/65, 10.8%; $\chi^2 = 16.836$, $p < 0.001$). The survival analysis further demonstrated that high ALDH1 expression was associated with poorer overall survival in clinical TNBC patients (HR = 4.11, 95% CI: 1.70–9.91, $p < 0.001$) (Table 2, Fig. 4).

Table 2 Association between ALDH1 protein expression and clinicopathological parameters ($N = 96$)

Clinicopathological Characteristic	Total N (%)	ALDH1 Expression		χ^2	*p*-value
		High (H-score ≥ 95) $n = 31$(%)	Low (H-score < 95) $n = 65$(%)		
Age (years)					
≤ 50	29 (30.2)	9 (29.0)	20 (30.8)	0.028	0.867
> 50	67 (69.8)	22 (71.0)	45 (69.2)		
Menopausal Status					
Postmenopausal	67 (69.8)	23 (74.2)	44 (67.7)	0.436	0.509
Premenopausal	29 (30.2)	8 (25.8)	21 (32.3)		
TNM Stage					
I	29 (30.2)	7 (22.6)	22 (33.8)	1.271	0.530
II	48 (50.0)	17 (54.8)	31 (47.7)		
III + IV	19 (19.8)	7 (22.6)	12 (18.5)		
Histological Grade					
Low-grade (G1/G2)	17 (17.7)	6 (19.4)	11 (16.9)	0.085	0.771
High-grade (G3)	79 (82.3)	25 (80.6)	54 (83.1)		
Tumor Size (cm)					
≤ 2	39 (40.6)	14 (45.2)	25 (38.5)	0.061	0.805
> 2	57 (59.4)	17 (54.8)	40 (61.5)		
Axillary Lymph Node Metastasis					
Positive	39 (40.6)	14 (45.2)	25 (38.5)	0.393	0.531
Negative	57 (59.4)	17 (54.8)	40 (61.5)		
Histological Type					
NOS	66 (68.8)	21 (67.7)	45 (69.2)	0.021	0.884
Special Types	30 (31.2)	10 (32.3)	20 (30.8)		
Ki-67 Index (%)					
≤ 20%	13 (13.5)	5 (16.1)	8 (12.3)	0.260	0.610
> 20%	83 (86.5)	26 (83.9)	57 (87.7)		
Survival Status					
Alive	74 (77.1)	16 (51.6)	58 (89.2)	16.836	< 0.001
Deceased	22 (22.9)	15 (48.4)	7 (10.8)		

[a]The percentages in the 'ALDH1 expression' columns (high and low) represent the proportion of each ALDH1 expression group (i.e., $n = 31$ for high, $n = 65$ for low)

[b]The percentages in the 'Total n (%)' column represent the proportion of the entire cohort ($n = 96$)

[c]Pearson's chi-square χ^2 test was used to assess associations between ALDH1 expression (categorical) and clinicopathological parameters (categorical). For TNM stage (ordinal variable with > 2 groups), the χ^2 test compared distributions across all stages

[d]NOS: Not Otherwise Specified

Discussion

TNBC manifests profound molecular heterogeneity and clinical aggressiveness given the paucity of effective targeted therapies. Against this backdrop, identifying validated biomarkers is critical for risk stratification and therapeutic development. Our integrated GEO analysis revealed significantly elevated ALDH1 mRNA in TNBC versus paired normal tissues ($p < 0.010$). Using clinical specimens, IHC validation substantiated marked ALDH1 overexpression in tumor lesions (H-score ≥ 95), which was associated with reduced survival (McNemar test, $p < 0.001$). This mortality risk pattern paralleled the Kaplan–Meier plotter data, where ALDH1-high patients exhibited substantially increased mortality (HR = 2.91, 95% CI: 1.21–6.96). The limited sample size likely contributed to the wide confidence interval. In our institutional cohort, ALDH1 overexpression correlated with reduced survival in TNBC patients. As a cancer stem cell marker, ALDH1 may reflect the intrinsic aggressiveness rather than the anatomical progression of a tumor, which accounts for its dissociation from conventional staging parameters. Its lack of association with conventional clinicopathological parameters in univariate analysis, combined with its prognostic value, suggests that ALDH1 may serve as a complementary biomarker to refine existing risk stratification frameworks.. Although inconsistencies exist in the literature regarding the prognostic value of ALDH1 in TNBC, potentially due to isoform-specific biological differences, methodological variations, and cohort heterogeneity [28–32]. We used an ALDH1A1 antibody. This could partially explain why our protein-level IHC results differ from some mRNA-based studies [31]. Some studies used a binary scoring system (positive/negative), whereas our study employed a more granular, continuous H-score method with a data-driven cut-off [32]. This methodological difference alone can greatly influence the patient stratification and subsequent statistical associations. Differences in cohort size, ethnic background, treatment regimens, and length of follow-up can all impact survival analysis outcomes. Our single-institution cohort ($n = 96$) may have different characteristics compared to other studies. Further validation should use isoform-specific assays and prospectively defined scoring criteria is needed to clarify the clinical utility of ALDH1 as a prognostic biomarker.

Extensive studies have established that high ALDH1 expression is both more prevalent and more robustly associated with poor prognosis in basal-like/TNBC subtypes than in hormone receptor-positive breast cancer [24, 33, 34]. Pharmacological studies utilizing TNBC xenograft models by Bousquet et al. revealed that chemotherapy-induced hypoxic conditions promote CSCs autophagy,

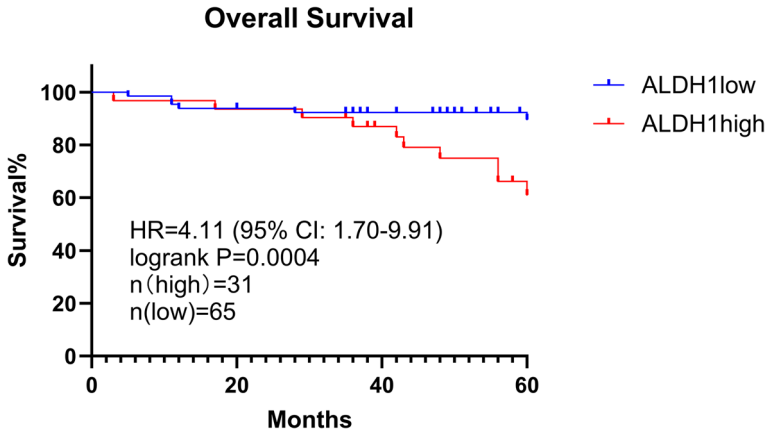

Fig. 4 Overall survival curves for patients with ALDH1 expression ($p < 0.001$). Analysis of clinical outcomes revealed that high ALDH1 expression (H-score ≥ 95, $n = 31$) was significantly associated with poorer survival compared to low ALDH1 expression (H-score < 95, $n = 65$) in TNBC patients (HR = 4.11, 95% CI: 1.70–9.91, log-rank $P < 0.001$)

thereby reducing therapeutic vulnerability to cytotoxic agents [35]. Experimental evidence indicates that targeted inhibition of these molecular mediators or microenvironmental modulation significantly decreases CSCs populations in TNBC, effectively reversing chemoresistance and suppressing metastatic progression [35, 36]. Furthermore, ALDH1 + TNBC cells exhibit increased resistance to chemotherapy and radiation, suggesting its role in maintaining CSC populations and driving treatment failure [37]. Ding et al. highlighted the unique immunosuppressive microenvironment of TNBC, characterized by enriched regulatory T cells (Tregs) and exhausted CD8 + T cells, which may interact with ALDH1 + CSCs to fosterpromote immune evasion [38]. Nevertheless, its functional significance within the CSC-enriched, treatment-resistant TNBC microenvironment remains poorly defined [37]. Although ALDH1 correlates with EMT progression and oxidative stress responses [37, 39], our findings raise a fundamental mechanistic question: whether ALDH1 overexpression actively drives TNBC pathogenesis or simply marks pre-existing CSC reservoirs. Paradoxically, emerging evidence suggests that ALDH1 may exert context-dependent tumor-suppressive roles, necessitating TNBC-specific validation via patient-derived organoids or relevant models.

Several study limitations warrant acknowledgment. The immunohistochemical cohort, while statistically robust, derives from a single institution, necessitating external validation through multi-center studies to confirm generalizability. Although bioinformatic analyses leveraged publicly available datasets, mechanistic validation of ALDH1-associated pathways—including retinoic acid metabolism and reactive oxygen species detoxification—remains to be established. Our immunohistochemical analysis utilized an antibody specific for ALDH1A1, whereas our mRNA data focused on the expression of ALDH1L2 and ALDH1B1. While all belong to the ALDH1 family, these isoforms exhibit distinct subcellular localizations and potentially different biological functions. Therefore, the protein and transcript readouts are not directly comparable as they measure different targets. Future studies are warranted to validate these findings at the protein level using antibodies specifically validated for ALDH1L2 and ALDH1B1. This would provide a more comprehensive understanding of the role of these mitochondrial isoforms in TNBC pathogenesis and stemness. The exclusion of 44 cases due to the absence of adjacent benign tissue may have biased our cohort toward patients with less advanced disease. Its limit the generalizability of our survival findings. This study demonstrates that elevated ALDH1 expression in TNBC tissues correlates with reduced overall survival ($p < 0.001$) and higher mortality risk, without significant associations to conventional clinicopathological parameters. While these findings support ALDH1 as a potential prognostic biomarker, its clinical utility requires further validation in multivariate-adjusted analyses. Future research should implement integrated experimental frameworks combining patient-derived organoids with genetically engineered mouse models to delineate two fundamental aspects of ALDH1 biology: (i) its regulatory crosstalk with immune checkpoint machinery and core stemness transcription factors (e.g., OCT4, SOX2, NANOG); (ii) the development of standardized immunohistochemical scoring criteria to facilitate clinical translation of ALDH1 as a prognostic biomarker.

Conclusion

This study establishes ALDH1 as a potential biomarker in TNBC, with elevated expression in malignant versus normal tissues (mRNA: $p < 0.01$; protein: H-score ≥ 95, McNemar $p < 0.001$) robustly correlating with reduced overall survival (HR = 4.11, 95% CI: 1.70–9.91). The integration of ALDH1 assessment into clinical stratification frameworks may enhance prognostication and guide targeted therapeutic strategies for high-risk TNBC patients.

Abbreviations

TNBC	Triple-negative breast cancer
ALDH1	Aldehyde dehydrogenase 1
IHC	Immunohistochemistry
TMA	Tissue microarray
GEO	Gene expression omnibus
ER	Estrogen receptor
PR	Progesterone receptor
HER2	Human epidermal growth factor receptor 2
CSCs	Cancer stem cells
EMT	Epithelial-mesenchymal transition
NF-κB	Nuclear factor-kappa B
IL-6	Interleukin-6
IL-8	Interleukin-8
ROS	Reactive oxygen species
RA	Retinoic acid
FFPE	Formalin-fixed paraffin-embedded
WHO	World health organization
HIER	Heat-induced epitope retrieval
H-score	Histochemical score
H&E	Hematoxylin and eosin
OCT4	Octamer-binding transcription factor 4
SOX2	SRY-box transcription factor 2
NANOG	Homeobox protein NANOG
NOS	Not otherwise specified

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13000-025-01726-y.

Additional file 1.

Additional file 2.

Acknowledgements

We would like to express our gratitude to all individuals who have provided valuable suggestions and comments, contributing to the enhancement of our paper's quality. We extend our sincere appreciation for the unwavering support provided by all the funding agencies for this study.

Authors' contributions

YY Y: Conceptualization, Methodology, Data curation, Writing- Original draft. ZR L: Data curation, Writing—original draft, Writing—review and editing, Visualization. YY Z: Data curation, Validation, Investigation. CY G: Formal analysis, Writing—Reviewing and Editing. L Y: Resources, Supervision. J Z: Funding acquisition, Project administration.

Funding

This work was supported by the Shandong Provincial Medical and Health Science and Technology Development Program (202101040677, 202304080192, 2018WS569).

Data availability

Publicly available datasets (GSE38959, GSE52194) were analyzed in this study. These data can be found in the GEO repository: https://www.ncbi.nlm.nih.gov/geo/

Declarations

Ethics approval and consent to participate

Written informed consent was obtained prior to tissue collection.

Competing interests

The authors declare no competing interests.

Received: 6 August 2025 / Accepted: 6 October 2025
Published online: 16 October 2025

References

1. Pajewska M, et al. Advanced and metastatic triple negative breast cancer—potential new treatment. Cancers. 2025;17:1183.
2. Bray F, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74:229–63.
3. Lin NU, et al. Clinicopathological features, patterns of recurrence, and survival among women with triple-negative breast cancer in the National Comprehensive Cancer Network. Cancer. 2012;118:5463–72.
4. Li Y, et al. Recent advances in therapeutic strategies for triple-negative breast cancer. J Hematol Oncol. 2022;15:121.
5. Mehanna J, Haddad FG, Eid R, Lambertini M, Kourie HR. Triple-negative breast cancer: current perspective on the evolving therapeutic landscape. Int J Womens Health. 2019;11:431–7.
6. Xia M, Zu X, Chen Z, Wen G, Zhong J. Noncoding RNAs in triple negative breast cancer: mechanisms for chemoresistance. Cancer Lett. 2021;523:100–10.
7. Bianchini G, De Angelis C, Licata L, Gianni L. Treatment landscape of triple-negative breast cancer - expanded options, evolving needs. Nat Rev Clin Oncol. 2022;19:91–113.
8. Vicente-Dueñas C, Romero-Camarero I, Cobaleda C, Sánchez-García I. Function of oncogenes in cancer development: a changing paradigm. EMBO J. 2013;32:1502–13.
9. Brooks MD, Wicha MS. Tumor twitter: cellular communication in the breast cancer stem cell niche. Cancer Discov. 2015;5:469–71.
10. Liu S, et al. Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. Stem Cell Rep. 2013;2:78–91.
11. Luo M, Wicha MS. Targeting cancer stem cell redox metabolism to enhance therapy responses. Semin Radiat Oncol. 2019;29:42–54.
12. Pai SG, et al. Wnt/beta-catenin pathway: modulating anticancer immune response. J Hematol Oncol. 2017;10(1):101.
13. Lu H, et al. A breast cancer stem cell niche supported by juxtacrine signaling from monocytes and macrophages. Nat Cell Biol. 2014;16:1105–17.
14. Shao C, et al. Essential role of aldehyde dehydrogenase 1A3 (ALDH1A3) for the maintenance of non-small cell lung cancer stem cells is associated with the STAT3 pathway. Clin Cancer Res Off J Am Assoc Cancer Res. 2014;20:4154–66.
15. Luo Y, et al. ALDH1A isozymes are markers of human melanoma stem cells and potential therapeutic targets. Stem Cells. 2012;30:2100–13.
16. Cheng P, et al. FOXD1-ALDH1A3 signaling is a determinant for the self-renewal and tumorigenicity of mesenchymal glioma stem cells. Cancer Res. 2016;76:7219–30.
17. Chen M-H, et al. ALDH1A3, the major aldehyde dehydrogenase isoform in human cholangiocarcinoma cells, affects prognosis and gemcitabine resistance in cholangiocarcinoma patients. Clin Cancer Res. 2016;22:4225–35.
18. Sullivan KE, Rojas K, Cerione RA, Nakano I, Wilson KF. The stem cell/cancer stem cell marker ALDH1A3 regulates the expression of the survival factor tissue transglutaminase, in mesenchymal glioma stem cells. Oncotarget. 2017;8:22325–43.
19. Xu X, et al. A self-accelerating 'copper bomb' strategy activated innate and adaptive immune response against triple-negative breast cancer. Bioact Mater. 2025;49:193–206.
20. Lestari DY, Mastutik G, Mukono IS. Betulinic acid and oleanolic acid modulate CD81 expression and induce apoptosis in triple-negative breast cancer cells through ROS generation. Med Oncol. 2024;42:24.
21. Li G, et al. Glycometabolic reprogramming-induced XRCC1 lactylation confers therapeutic resistance in ALDH1A3-overexpressing glioblastoma. Cell Metab. 2024;36:1696-1710.e10.
22. Marcato P, et al. Aldehyde dehydrogenase 1A3 influences breast cancer progression via differential retinoic acid signaling. Mol Oncol. 2015;9:17–31.
23. Bu J, et al. Kk-lc-1 as a therapeutic target to eliminate ALDH+ stem cells in triple negative breast cancer. Nat Commun. 2023;14:2602.
24. Ginestier C, et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell. 2007;1:555–67.
25. Nagy Á, Lánczky A, Menyhárt O, Győrffy B. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. Sci Rep. 2018;8:9227.
26. Zhang J, Liu S, Wu M, Shi W, Cai Y. Clinical significance and expression of SLC35F6 in bladder urothelial carcinoma. Diagn Pathol. 2024;19:150.

27. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clin Cancer Res. 2004;10:7252–9.
28. Ohi Y, et al. Aldehyde dehydrogenase 1 expression predicts poor prognosis in triple-negative breast cancer. Histopathology. 2011;59:776–80.
29. Ma F, et al. Aldehyde dehydrogenase 1 (ALDH1) expression is an independent prognostic factor in triple negative breast cancer (TNBC). Medicine (Baltimore). 2017;96:e6561.
30. Yao J, Jin Q, Wang X-D, Zhu H-J, Ni Q-C. Aldehyde dehydrogenase 1 expression is correlated with poor prognosis in breast cancer. Medicine (Baltimore). 2017;96:e7171.
31. Liu Y, et al. ALDH1A1 mRNA expression in association with prognosis of triple-negative breast cancer. Oncotarget. 2015;6:41360–9.
32. Demir H, Dulgar O, Gulle BT, Turna H, Ilvan S. Prognostic value of aldehyde dehydrogenase 1 (ALDH1) in invasive breast carcinomas. Bosn J Basic Med Sci. 2018;18:313–9.
33. Ginestier C, et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell. 2007;1:555–67.
34. Ricardo S, et al. Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype. J Clin Pathol. 2011;64:937–46.
35. Bousquet G, et al. Targeting autophagic cancer stem-cells to reverse chemoresistance in human triple negative breast cancer. Oncotarget. 2017;8:35205–21.
36. Cazet AS, et al. Targeting stromal remodeling and cancer stem cell plasticity overcomes chemoresistance in triple negative breast cancer. Nat Commun. 2018;9:2897.
37. Bu J, et al. Kk-lc-1 as a therapeutic target to eliminate ALDH+ stem cells in triple negative breast cancer. Nat Commun. 2023;14:2602.
38. Ding S, et al. Single-cell atlas reveals a distinct immune profile fostered by T cell-B cell crosstalk in triple negative breast cancer. Cancer Commun. 2023;43:661–84.
39. Zheng Y, et al. A diagnostic model for sepsis-induced acute lung injury using a consensus machine learning approach and its therapeutic implications. J Transl Med. 2023;21:620.

Publisher's note