Received: 24 July 2023 | Revised: 12 January 2024 | Accepted: 29 February 2024

DOI: 10.1002/cac2.12530

CANCER COMMUNICATIONS Open Access

LETTER TO THE JOURNAL

Discovering genetic biomarkers for targeted cancer therapeutics with eXplainable Artificial Intelligence

High-Grade Serous Ovarian Cancer (HGSC) is the most prevalent and lethal form of gynecologic malignancies [1], accounting for 70%-80% of ovarian cancer fatalities. Despite decades of research, the overall survival rate for HGSC has remained largely unchanged [2], and patients with advanced stages of the disease have only a 41% chance of surviving beyond five years [3]. Investigating the genomic and immune profiles of long-term HGSC survivors could offer valuable insights into the underlying tumor biology and inform potential therapeutic strategies [4]. This study advances upon prior research by employing an innovative eXplainable Artificial Intelligence (XAI) integrated with a hypothesis-driven probabilistic methodology to dissect the intricate genetic underpinnings linked to HGSC's survival outcomes in a cohort of 407 patients. The objective of this article was to uncover the most critical prognostic biomarkers from a pool of 655 potential targets through our distinctive data-driven approach and determine the impacts of potentially modulating the identified biomarkers on HGSC outcomes.

Recent studies indicate that AI models are often referred to as "black boxes" because their decision-making process lacks transparency [5]. The consensus is that the lack of inherent explainability is problematic as this produces biases, creates difficulties in detecting false positives and negatives, and conceals potential insights that may be derived from AI [6]. In this study, we provide evidence demonstrating how XAI can enhance biological explainability by revealing novel insights from the underlying data (Supplementary Materials and Methods). Our XAI approach distinctively predicts patient outcomes and survival duration based on genetic signatures (predictive AI aspect of the models) and discovers and helps visualize critical biomarkers (biological explainability aspect of the models) in HGSC. To ensure the viability of explanations generated by our XAI, we subsequently validated the most prominent HGSC-promoting biomarker identified by XAI using in vivo murine tumor models (Supplementary Figure S1). The XAI approach outlined in this study is a proof-of-concept that is not only intended to generate high predictive accuracy but also infer the cause-effect relations behind the predictions, identify counterfactuals that are useful for optimizing interventional therapies, and assess the resultant improvements in patients.

We report that our models predicted the ≥5-year overall survival probability based on the genetic features of patients (n = 407) with 97.52% accuracy, 100% precision, and 94.74% recall on the testing data that comprised 25% of the total samples that were hidden from the models during the training phase. Insights derived through XAI prioritized the biomarkers that are of utmost importance in determining prognosis for patients with HGSC, which we refer to as 'global biological explanations' (Figure 1A). The biomarkers are arranged in terms of their relative importance (most important at the top), and the variations in the expression of the biomarkers from high (red) to low (blue) together with the corresponding Shapley values on the x-axis were used to determine whether a particular biomarker is associated with poor (positive values on the x-axis) or good prognosis (negative values on the x-axis) in HGSC (Figure 1A). For example, we discovered that TATA-Box Binding Protein Associated Factor 10 (TAF10) is associated with good prognosis since higher concentrations (red) of this biomarker correspond to negative Shapley values on the x-axis; hence, it can be classified as a cancer suppressor for HGSC patients. In contrast, Interleukin 27 receptor subunit alpha (IL27RA) is associated with poor prognosis since higher concentrations (red) of this biomarker correspond to positive Shapley values on the x-axis; hence, it can be classified as a cancer promoter for HGSC patients. Furthermore, XAI unveiled the inflection points of each influential gene (local biological explanations) above or below which the prognosis would either improve or deteriorate (Figure 1B). These inflection

List of abbreviations: AI, artificial intelligence; DOPC, 1,2-Dioleoyl-sn-glycero-3-phosphocholine; HGSC, high-grade serous ovarian cancer; IL27RA, Interleukin 27 receptor subunit alpha; siRNA, silencing RNA; TAF10, TATA-Box Binding Protein Associated Factor 10; TPM, transcripts per million; XAI, eXplainable Artificial Intelligence.

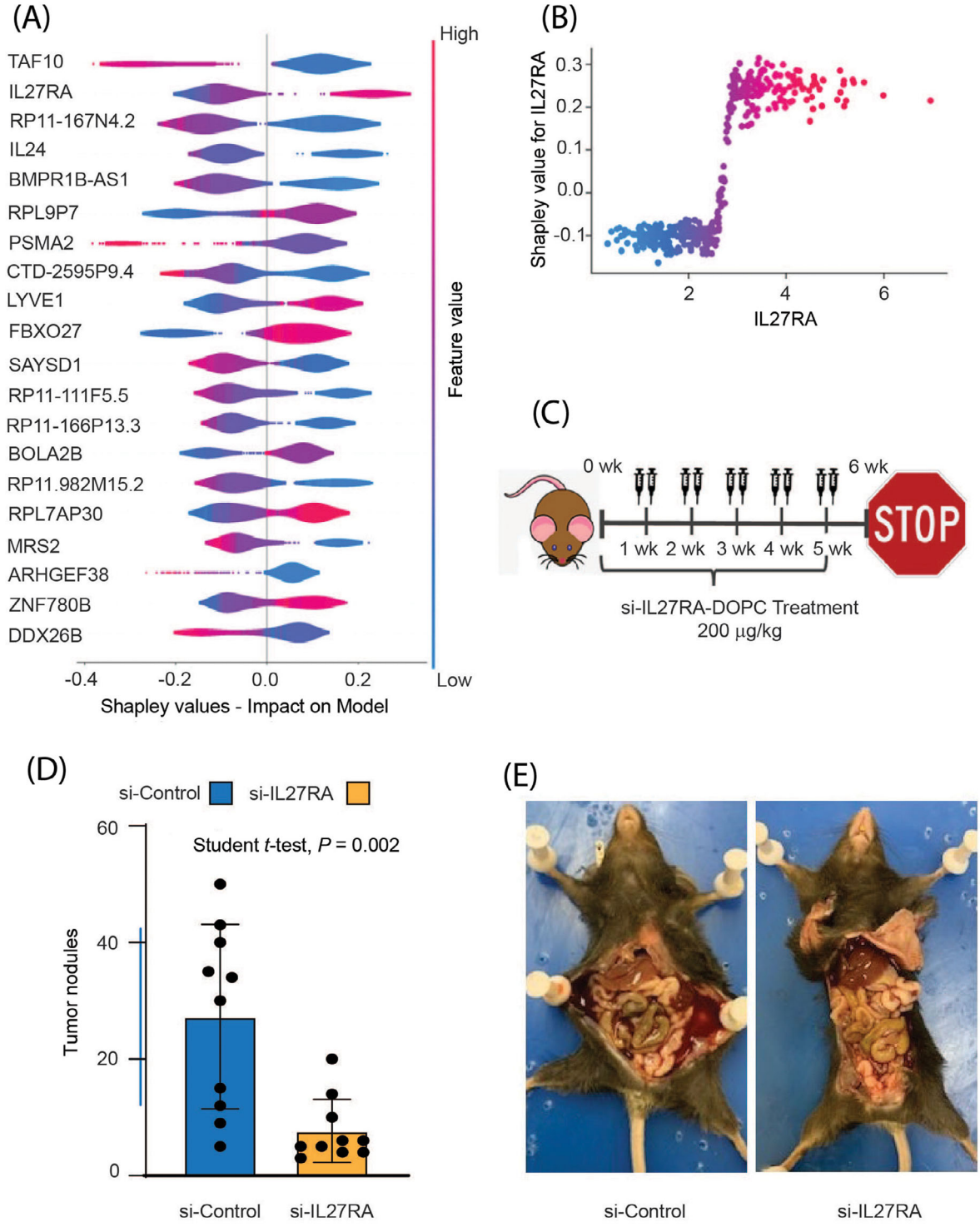

FIGURE 1 Comprehensive illustrations of XAI-derived findings and in vivo experimental results in HGSC research. (A) Global biological explanations: This panel presents a comprehensive visual summary of the critical biomarkers identified by the XAI approach in HGSC. It displays a ranked list of biomarkers based on their relative importance in determining patient prognosis, with variations in biomarker expression levels depicted through a color gradient from high (red) to low (blue). Accompanying Shapley values on the x-axis provide insights into whether a biomarker is associated with positive (cancer-promoting) or negative (cancer-suppressing) outcomes in HGSC. (B) Local biological explanations for IL27RA: This section focuses specifically on the IL27RA gene, identified as a significant cancer-promoting biomarker. It offers a detailed analysis of how variations in IL27RA expression influence HGSC prognosis. The visualization includes inflection points indicating the expression levels above or below which the prognosis for patients may improve or deteriorate. (C) si-IL27RA treatment schedule in in vivo model: This panel outlines the experimental protocol used in the in vivo study, detailing the schedule and dosage of siRNA (silencing RNA) targeting IL27RA administered to C57BL/6 mice. (D) Statistical comparison of

points serve as guidance for genomic editing that could lead to the development of new gene therapeutics and enhanced prognosis of patients.

Drawing on the insights from XAI, our research identified IL27RA as a key biomarker linked to lower survival rates in HGSC patients. The XAI results led us to theorize that reducing IL27RA levels to below 2.2 transcripts per million (TPM) might improve patient outcomes (Figure 1B). To test this hypothesis, we employed an orthotopic ovarian cancer model in C57BL/6 mice, which were intraperitoneally inoculated with one million ID8 ovarian cancer cells. This model closely mirrors the extensive intraperitoneal spread observed in the human form of the disease. We divided twenty mice into two groups of ten each. Both groups were intraperitoneally injected with tumor cells. The first group received intravenous injections of 1,2-Dioleoyl-sn-glycero-3-phosphocholine (DOPC) nanoparticles containing scrambled siRNA, while the second group received DOPC nanoparticles loaded with IL27-siRNA (si-IL27RA-DOPC treatment) (Figure 1C). Over five weeks, these treatments were administered bi-weekly. Post-treatment, the mice were euthanized, and their peritoneal metastases were quantified. The si-IL27RA-DOPC treatment group exhibited a significant decrease in peritoneal metastasis (Figure 1D-E). These results indicate that targeting IL27RA may be a promising approach to improving prognoses in HGSC patients.

The biology of IL27 is multifaceted and varies depending on the context, exhibiting the potential to either mitigate or exacerbate various forms of inflammation and cancer [7]. This complexity underscores its dual role in immune modulation across different biological scenarios. Recently, an experimental study revealed that IL27 receptor signaling promoted hepatocellular carcinoma development in vivo and that its high expression correlated with poor prognosis for patients due to increased proliferative capacity of tumors and inflammation [8]. Thus, further validating that the explanations obtained from XAI could be in fact legitimate. We adopted a probabilistic approach to evaluate insights unraveled from the XAI models relating to the underlying relationships between the features and targets (Supplementary Figure S2). Our study revealed that strategic modulation of the top five biomarkers in combination (as illustrated in Figure 1A and Supplementary Figure S2) could potentially enhance the survival prospects of patients with HGSC by up to 100% for periods extending beyond five years. Notably, TAF10 emerged as a prominent HGSC suppressor according to our XAI analysis (Figure 1A and Supplementary Figures S2A and S2F), with its individual modulation potentially contributing to a 41.1% increase in survival likelihood. Further supporting these findings, Kaplan-Meier and box-plot analyses of TAF10 and IL27RA aligned with the predictions of our XAI models (Supplementary Figure S3). These analyses demonstrated a positive correlation between higher TAF10 expression and increased survival chances. Conversely, an elevated expression of IL27RA was linked to diminished survival prospects, as detailed in Supplementary Figure S3. Based on the evidence from this study, XAI approaches are amenable to generating biologically relevant testable hypotheses despite their limitations due to explanations originating from post hoc realizations [9]. Additionally, the application of probabilistic causality methods can enhance the insights gained from XAI. This synergistic use of both approaches offers a more streamlined pathway to propel advancements in precision oncology, yielding time and cost efficiencies.

In conclusion, this study presents a groundbreaking integrated approach combining XAI and probabilistic methods to advance the understanding and treatment of HGSC. Our research focused on deciphering the complex interactions between genetic biomarkers and the five-year survival probabilities of HGSC patients, utilizing a data set of 407 samples. The unique blend of XAI and probabilistic causality approaches has not only demonstrated high predictive accuracy but also provided valuable biological explanations, identifying key biomarkers that influence patient outcomes. Particularly, our findings highlight the roles of biomarkers like TAF10 and IL27RA in affecting survival rates, revealing their potential as therapeutic targets. Furthermore, our in vivo experiments using murine tumor models have validated the XAI-derived hypotheses, particularly regarding the role of IL27RA in HGSC prognosis. This study not only enhances our understanding of HGSC but also illustrates the potential of XAI in precision oncology, offering a more efficient pathway for developing targeted therapies. As we continue to explore and refine these approaches, the prospects for their application in diverse cancer types and clinical settings are promising, paving the way for more personalized and effective cancer treatments.

metastasis: A comparison of metastatic spread, represented by the count of tumor nodules, between the control group and the group treated with si-IL27RA. This graphically represented data provides a clear and quantifiable measure of the effectiveness of the si-IL27RA treatment in reducing tumor metastasis. (E) Representative images of tumor-bearing mice: This section includes photographic evidence from the in vivo study, showing C57BL/6 mice bearing ID8 tumors. These images serve as a visual representation of the physical outcomes of the experiment, comparing the control and treated groups, and providing a tangible aspect to the study's findings.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: Debaditya Chakraborty, Hakan Başağaoğlu and Gabriel Lopez-Berestein. *Performed the experiments*: Debaditya Chakraborty, Paola Amero and Cristian Rodriguez-Aguayo. *Analyzed the data*: All authors. *Contributed materials/analysis tools*: All authors. *Wrote the paper*: Debaditya Chakraborty and Elizabeth Gutierrez-Chakraborty primarily and all authors reviewed, revised, and commented on the manuscript.

ACKNOWLEDGMENTS

We thank Dr. Cristina Ivan for helping with data acquisition while at MDACC.

CONFLICT OF INTEREST STATEMENT

The authors declare that there are no competing interests.

FUNDING INFORMATION

This study was supported in part by grants from the National Institutes of Health/National Cancer Institute (5U01CA213759-02, P30CA016672), and the American Cancer Society, National Science Foundation (CHE-1411859), and an endowment grant from the John P. Gaines Foundation. Dr. Cristian Rodriguez-Aguayo and Dr. Paola Amero were supported by the Brain SPORE Career Enhancement Program and NCI grant P50CA127001, as well as by the NIH through the Ovarian SPORE Career Enhancement Program and NCI grant P50CA217685.

DATA AVAILABILITY STATEMENT

The datasets and code for this study are available upon request.

ETHICS APPROVAL

All mouse studies were approved by the Institutional Animal Care and Use Committee at MD Anderson. All animal experiments were performed with 4- to 6-week-old female C57BL/6 mice obtained from Taconic Biosciences (Rensselaer, NY). Five mice per cage were housed under pathogen-free conditions at a constant temperature and humidity. All mice were fed a regular diet and water ad libitum according to American Association for Laboratory Animal Science guidelines and the US Public Health Service Policy on Human Care and Use of Laboratory Animals. Mice were euthanized via cervical dislocation if found moribund by the investigators, who have approval from the IACUC committee (IACUC Study #00001010-RN03).

Debaditya Chakraborty[1]
Elizabeth Gutierrez-Chakraborty[2]
Cristian Rodriguez-Aguayo[3]
Hakan Başağaoğlu[4]
Gabriel Lopez-Berestein[3]
Paola Amero[3]

[1]*School of Civil and Environmental Engineering and Construction Management, The University of Texas at San Antonio, San Antonio, Texas, USA*
[2]*Department of Cell Systems and Anatomy, University of Texas Health Science Center, San Antonio, Texas, USA*
[3]*Department of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA*
[4]*Edwards Aquifer Authority, San Antonio, Texas, USA*

Correspondence
Debaditya Chakraborty, School of Civil and Environmental Engineering and Construction Management, The University of Texas at San Antonio, San Antonio, Texas, USA.
Email: debaditya.chakraborty@utsa.edu

Paola Amero, Department of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA.
Email: PAmero@mdanderson.org

ORCID

Debaditya Chakraborty https://orcid.org/0000-0002-2165-9440

REFERENCES

1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. CA: A Cancer Journal for Clinicians. 2023;73(1):17-48.
2. Bowtell DD, Böhm S, Ahmed AA, Aspuria PJ, Bast RC, Jr., Beral V, et al. Rethinking ovarian cancer II: reducing mortality from high-grade serous ovarian cancer. Nat Rev Cancer. 2015;15(11):668-679.
3. Millstein J, Budden T, Goode EL, Anglesio MS, Talhouk A, Intermaggio MP, et al. Prognostic gene expression signature for high-grade serous ovarian cancer. Ann Oncol. 2020;31(9):1240-1250.
4. Garsed DW, Pandey A, Fereday S, Kennedy CJ, Takahashi K, Alsop K, et al. The genomic and immune landscape of long-term survivors of high-grade serous ovarian cancer. Nature genetics. 2022;54(12):1853-1864.
5. Shmatko A, Ghaffari Laleh N, Gerstung M, Kather JN. Artificial intelligence in histopathology: enhancing cancer research and clinical oncology. Nat Cancer. 2022;3(9):1026-1038.
6. Kundu S. AI in medicine must be explainable. Nature Medicine. 2021;27(8):1328.
7. Dibra D, Mitra A, Newman M, Xia X, Cutrera JJ, Gagea M, et al. Lack of Immunomodulatory Interleukin-27 Enhances Oncogenic Properties of Mutant p53 In Vivo. Clin Cancer Res. 2016;22(15):3876-3883.
8. Aghayev T, Mazitova AM, Fang JR, Peshkova IO, Rausch M, Hung M, et al. IL27 Signaling Serves as an Immunologic

Checkpoint for Innate Cytotoxic Cells to Promote Hepatocellular Carcinoma. Cancer discovery. 2022;12(8):1960-1983.
9. Babic B, Gerke S, Evgeniou T, Cohen IG. Beware explanations from AI in health care. Science. 2021;373;284-286.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.