

BRIEF COMMUNICATION



Interpretable machine learning identifies health system levers for survival outcomes of patients with prostate cancer

Milit S. Patel^{1,2✉}, James Janopaul-Naylor^{1b}, Frederic Ivan L. Ting^{3,4}, James Fan Wu⁵, Sean M. McBride², Dana E. Rathkopf², Erin Jay G. Feliciano^{6,7}, Hong Truong⁸, Puneeth Iyengar^{1b}, Jonas Willmann^{9,10}, Daniel E. Gorovets², Paul L. Nguyen¹¹, Vedang Murthy^{1b}, Brandon A. Mahal^{1b}, Nancy Y. Lee², Himanshu Nagar^{1b} and Edward Christopher Dee^{1b}

© The Author(s), under exclusive licence to Springer Nature Limited 2026

BACKGROUND: Advances in the care of patients with prostate cancer across the disease spectrum have resulted in substantial survival gains over recent decades. However, global disparities in prostate cancer survival, stage at diagnosis, and access to timely diagnosis and treatment remain among the greatest across all malignancies. We aimed to quantify country-level health system correlates of the prostate cancer mortality-to-incidence ratio (MIR).

METHODS: We applied interpretable machine learning using CatBoost with SHAP (SHapley Additive exPlanations) to all 185 countries with prostate cancer MIR estimates in GLOBOCAN 2022. National age-standardized MIR estimates were linked to 11 health system indicators from the World Health Organization Global Health Observatory, World Bank World Development Indicators, United Nations agencies, and the Directory of Radiotherapy Centres. Model performance was evaluated using repeated leave-one-country-out cross-validation with bootstrap-based uncertainty estimation. CatBoost handled missing predictor values natively, allowing inclusion of all countries without imputation.

RESULTS: The model demonstrated strong predictive performance for prostate cancer MIR ($R^2 = 0.80$, $RMSE = 0.078$, $MAE = 0.060$, Pearson correlation = 0.89). The health system features most strongly associated with lower MIR were gross domestic product per capita, the universal health coverage (UHC) service coverage index, and radiotherapy infrastructure. Country-level SHAP decompositions demonstrated heterogeneous feature contributions across settings, identifying context-specific priorities. Greater radiotherapy capacity and more comprehensive UHC were consistently associated with lower MIR, whereas higher aggregate health expenditure alone showed weaker associations with improved outcomes.

CONCLUSIONS: Interpretable machine learning identified radiotherapy infrastructure, universal health coverage, and national wealth as the health system features most consistently associated with lower prostate cancer mortality-to-incidence ratios across countries. These findings suggest that targeted investment in treatment capacity and service coverage may have greater relevance to prostate cancer outcomes than increases in aggregate health spending alone. Although ecological and hypothesis-generating, this framework may inform future health systems research, prospective evaluation, and global prostate cancer control.

Prostate Cancer and Prostatic Diseases; <https://doi.org/10.1038/s41391-026-01156-x>

INTRODUCTION

Advances in the care of patients diagnosed with prostate cancer across the disease spectrum have resulted in immense survival gains over the past several decades. However, global disparities in prostate cancer survival, stage at diagnosis, and access to treatment are among the starkest across all cancer types [1]. Prostate cancer is one of the leading causes of cancer-related morbidity and mortality among men worldwide; while age-standardized incidence continues to rise in many regions, reductions in mortality have been more pronounced in countries

with advanced cancer care infrastructure, routine access to early detection, and multimodal curative-intent treatment [2]. The global burden of cancer, including prostate cancer, is expected to rise substantially through 2050, particularly in countries with limited resources for diagnosis and patient care, underscoring the urgent need for actionable, health system-oriented strategies to address persistent disparities in prostate cancer survival, stage at diagnosis, and access to treatment [3].

Thus, efforts are needed to better understand how health system factors may be associated with prostate cancer outcomes.

¹Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX, USA. ²Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, USA. ³Section of Medical Oncology, Department of Internal Medicine, Corazon Locsin Montelibano Memorial Regional Hospital, Bacolod, Philippines. ⁴Department of Clinical Sciences, College of Medicine, University of St. La Salle, Bacolod, Philippines. ⁵Division of Hematology and Oncology, Department of Medicine, Medical College of Wisconsin, Milwaukee, WI, USA. ⁶School of Medicine and Public Health, Ateneo de Manila University, Pasig City, Philippines. ⁷Department of Hematology/Oncology, Northwell Health Cancer Institute, New Hyde Park, NY, USA. ⁸Department of Surgery (Urology Service), Memorial Sloan Kettering Cancer Center, New York, NY, USA. ⁹Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, NY, USA. ¹⁰Department of Radiation Oncology, University of Zurich, Zurich, Switzerland. ¹¹Department of Radiation Oncology, Mass General Brigham, Boston, MA, USA. ¹²Department of Radiation Oncology, ACTREC, Tata Memorial Centre, Homi Bhabha National Institute, Mumbai, India. ¹³Department of Radiation Oncology, Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, Miami, FL, USA. ✉email: militpatel@utexas.edu; deee1@mskcc.org

Received: 7 March 2026 Revised: 1 June 2026 Accepted: 4 September 2026

Published online: 02 October 2026

Here, we applied explainable artificial intelligence (AI) to provide an interpretable, country-level attribution of the health system features most strongly associated with prostate cancer mortality-to-incidence ratios. Our aim is to generate ecological, hypothesis-generating signals that may help focus future policy analyses and targeted health system investment, rather than to estimate causal policy effects.

METHODS

We used the prostate cancer mortality-to-incidence ratio (MIR) as a pragmatic ecological indicator of cancer control performance, enabling benchmarking across diverse socioeconomic contexts. Although metrics such as 5-year relative survival and net survival are generally preferred for direct cross-country comparison of cancer burden, they are not reliably available for prostate cancer across most low- and middle-income countries because they require high-quality population-based cancer registries with long-term active follow-up, linkage to national life tables, and age-standardized expected survival estimates. By contrast, MIR can be derived from the age-standardized mortality and incidence estimates reported for virtually every country in GLOBOCAN 2022, making it the most consistently available ecological proxy for global cross-country comparison of prostate cancer outcomes [4]. MIR provides an ecological indicator for cross-country comparisons but does not directly estimate patient survival or treatment effectiveness [4, 5].

National age-standardized MIR estimates for prostate cancer in men were derived from GLOBOCAN 2022, and were linked to health system indicators drawn from the World Health Organization Global Health Observatory, the World Bank World Development Indicators, United Nations agencies, and the Directory of Radiotherapy Centres (DIRAC). Predictor variables included GDP per capita (current US\$), the Universal Health Coverage (UHC) service coverage index, radiotherapy treatment centers per 1000 population (DIRAC), physicians per 1000 population, nurses and midwives per 1000 population, surgical workforce per 1000 population, presence of national pathology services, current health expenditure as a percentage of GDP, out-of-pocket expenditure as a percentage of current health expenditure, the Human Development Index (HDI), and the Gender Inequality Index.

GLOBOCAN 2022 reports prostate cancer MIR estimates for 185 countries, all of which were included in the final analytic sample ($n = 185$). Country-level completeness across the 11 health system predictors ranged from 77.8% (surgical workforce per 1000 population) to 100% (GDP per capita, radiotherapy centers per 1000 population, and presence of national pathology services), with most predictors at $\geq 90\%$ completeness; predictor-specific completeness is summarized in Supplementary Table 1. In the primary analysis, missing numerical predictor values were handled using CatBoost's native missing-value processing, and no country was excluded because of missing health systems predictors [6]. Within each training fold, prior to model fitting, we applied iterative variance inflation factor (VIF) selection (threshold 10), removing the most collinear feature one at a time on the training countries only and then refitting CatBoost on the retained features; this fold-internal selection was applied to limit redundancy among health system indicators (e.g., between GDP per capita, HDI, and UHC index) without leaking test-country information. To accommodate nonlinear, context-dependent relationships and the sparse, mixed-type nature of country-level health system data, we implemented a CatBoost gradient-boosted regression ensemble with ordered boosting to mitigate target leakage. Model performance was estimated using repeated leave-one-country-out (LOCO) cross-validation with 10 repeats using different random seeds, yielding 1850 country-fold models in total. For each repeat, every country was held out once as the test fold while the remaining 184 countries were used for training; out-of-fold predictions were aggregated per country across the 10 repeats. Hyperparameters (number of iterations, learning rate, tree depth, L2 leaf regularization, random strength, bagging temperature, per-level column subsampling, and border count) were tuned via Optuna's tree-structured Parzen estimator (TPE; an approximate Bayesian optimization) on inner GroupKFold splits of the training countries only, with early stopping on a held-out inner validation split, so that no test-country information was used during tuning.

Uncertainty intervals for R^2 , RMSE, correlation, and per-country predictions were derived from the empirical distribution of fold-level metrics and bootstrap resampling (2000 replicates) of the out-of-fold predictions, reported as 95% percentile intervals [6]. SHAP (SHapley Additive

exPlanations) values were computed for each country using TreeSHAP on every included model run, and both mean absolute SHAP values (global importance) and signed per-country SHAP values (directional contributions) were aggregated by averaging across all cross-validation runs in which a given country appeared in the test fold; only these included runs contributed to the reported SHAP summaries [7].

This framework supports country-level prediction of prostate cancer MIR and ecological characterization of the relative contribution of system-level features across heterogeneous national contexts. To assess robustness, pre-specified sensitivity analyses repeated the primary modeling stratified by World Bank income group (high-income vs. low- and middle-income countries) using complete-case filtering within each stratum to ensure stable model fitting; consistency of the leading SHAP predictors was compared across strata. Additional planned sensitivity analyses (registry-quality restriction and Cook's distance outlier exclusion) are deferred to future work given the modest analytic sample size.

RESULTS

All 185 countries with a GLOBOCAN 2022 prostate cancer MIR estimate were included in the final analytic sample, with CatBoost's native missingness handling applied to individual health system predictors. Out-of-fold predictions showed strong agreement with observed country-level MIR for prostate cancer (Fig. 1), with $R^2 = 0.796$ (95% CI: 0.732–0.844), RMSE = 0.078 (95% CI: 0.070–0.087), mean absolute error = 0.060 (95% CI: 0.053–0.068), and Pearson correlation = 0.892; complementary metrics (R^2 , RMSE, MAE, and correlation together) were used rather than RMSE alone to better characterize both bias and variance. A pre-specified sensitivity analysis stratified by World Bank income group (high-income, $n = 45$; low- and middle-income, $n = 79$; complete-case within each stratum) showed that the leading SHAP predictors (GDP per capita, radiotherapy center density, UHC index) remained consistent across strata; the corresponding scatter plots of observed versus predicted MIR are provided in Supplementary Fig. 1 (LMIC) and 2 (HIC). Additional planned sensitivity analyses (registry-quality restriction and Cook's distance outlier exclusion) are deferred to future work given the modest analytic sample size.

On the SHAP summary (Fig. 2), GDP per capita was the most influential predictor of country-level prostate cancer MIR, followed by radiotherapy treatment center density (DIRAC, per 1000 population) and the UHC service coverage index. Additional contributors, in descending order of mean absolute SHAP, were the Gender Inequality Index, physicians per 1000 population, current health expenditure as a percentage of GDP, nurses and midwives per 1000 population, out-of-pocket expenditure as a percentage of current health expenditure, surgical workforce per 1000 population, and presence of national pathology services. The income-group stratified sensitivity analysis described above (Supplementary Figs. 1, 2) showed that these leading SHAP predictors remained consistent across high-income and low- and middle-income strata.

DISCUSSION

These leading SHAP features (GDP per capita, radiotherapy center density, and the UHC service coverage index) are broadly consistent with prior pan-cancer analyses [4, 5] and highlight the roles of treatment availability and service coverage in shaping ecological MIR patterns for prostate cancer specifically. Country-level SHAP decompositions then showed that the most important features varied by setting. In higher-income countries, greater radiotherapy and health workforce capacity were most often associated with lower MIR, while in resource-constrained countries, lower UHC service coverage and weak treatment infrastructure more commonly appeared among the features contributing to higher predicted MIR. Previous linear model analyses identified associations between health systems indicators

Model Performance: Actual vs. Predicted MIR with Repeat-based Confidence Intervals

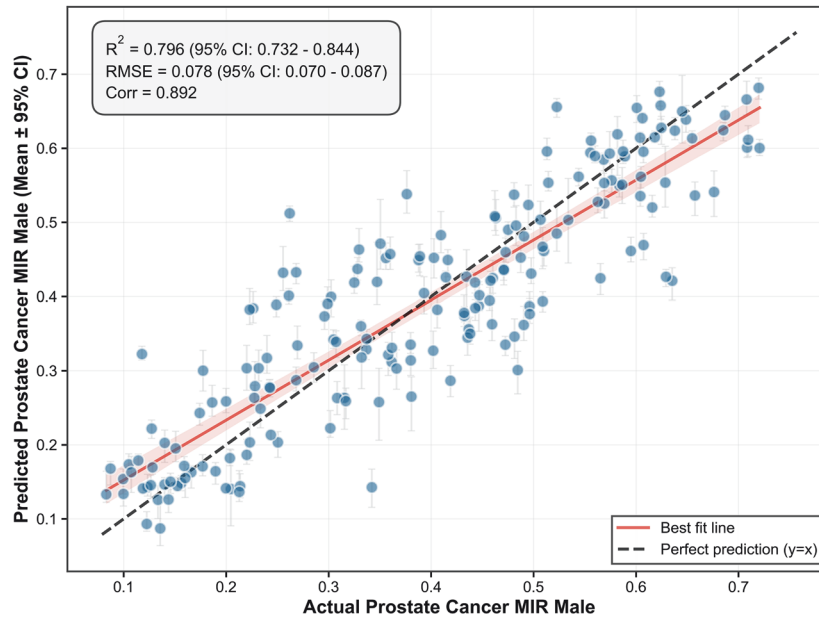


Fig. 1 Model performance: actual vs. predicted prostate cancer MIR in males. Scatter plot depicting the relationship between actual versus predicted prostate cancer mortality-to-incidence ratio (MIR) for males across countries. Each point represents a country's mean MIR with associated 95% confidence intervals. The model displayed strong predictive performance, with $R^2 = 0.796$ (95% CI: 0.732–0.844), RMSE of 0.078 (95% CI: 0.070–0.087), and correlation of 0.892. The dashed line indicates perfect prediction ($y = x$), while the solid regression line with shaded confidence band demonstrates the model fit.

and pan-cancer MIR [4]. Linear models, however, are more limited in their ability to handle missing predictor data and may disproportionately weight variables available mainly in higher-resource settings, potentially introducing selection bias. The SHAP-based CatBoost analysis better accommodates missingness and nonlinear effects, providing a more representative view of the country-level features associated with prostate cancer MIR. In prior ecological work using a related dataset, a higher ratio of social-to-medical health spending was modestly associated with lower cancer incidence broadly, but showed a weaker association with MIR [4, 5]. This is consistent with the present finding that aggregate health expenditure as a share of GDP carried less weight in the SHAP attribution than targeted infrastructure indicators (radiotherapy center density and UHC service coverage), suggesting that the composition and targeting of investment may matter more for prostate cancer MIR than aggregate spending levels alone.

As illustrative examples, Yemen and the United States occupied opposite ends of the observed MIR distribution. For Yemen (observed MIR 0.720, predicted 0.682), the country-level SHAP decomposition identified low GDP per capita (mean absolute SHAP contribution: 0.090) and high Gender Inequality Index (SHAP: 0.073) as the largest contributors pushing predicted MIR upward, followed by limited radiotherapy infrastructure (SHAP: 0.035) and low UHC service coverage (SHAP: 0.029). These are model-derived associations and should not be interpreted as causal estimates; external context, including Yemen's ongoing humanitarian crisis, healthcare system disruption, limited diagnostic capacity, and barriers to specialized oncology services, is consistent with but not directly inferred from the model output [8]. The United States, by contrast, had the low observed MIR (0.107; predicted 0.163).

For the U.S., the SHAP decomposition attributed the model's relatively low predicted MIR primarily to higher radiotherapy center density (SHAP: 0.043), higher current health expenditure as a percentage of GDP (SHAP: 0.035), and higher UHC service coverage (SHAP: 0.035) [9]. Contextually, and separate from the

model output, prior U.S. policy investments in Medicare coverage expansion, National Cancer Institute-designated comprehensive cancer centers, and broad adoption of intensity-modulated radiation therapy and advanced surgical techniques have contributed to infrastructure supporting early detection and definitive treatment [10]. The model's modest overprediction of U.S. MIR indicates that, within the feature set used here, the U.S. observed MIR is lower than the model would predict from its health system inputs alone. The causes of this discrepancy cannot be determined from the present ecological model.

Situated between these extremes, India's prostate cancer mortality-to-incidence ratio (0.49) reflects an intermediate outcome profile amid the global spectrum. Highlighting this, India's recent launch and rapid expansion of the Ayushman Bharat universal health coverage initiative stands as one of the largest efforts worldwide to enhance population-level financial protection and access to cancer care [11]. This program exemplifies the critical role of health system strengthening in middle-income countries seeking to improve cancer outcomes on a national scale.

The interpretability of SHAP allows country-level identification of features most strongly associated with predicted MIR. In several high-spending countries, for example, higher aggregate health spending was not accompanied by proportionally lower predicted MIR in the absence of higher radiotherapy capacity and UHC service coverage, consistent with the hypothesis that directed, infrastructure-oriented investment may be more relevant to prostate cancer MIR than aggregate spending alone. Such patterns, together with SHAP-informed national summaries, may help ministries of health and global partners such as WHO and the IAEA prioritize further evaluation of specific infrastructure gaps associated with higher MIR. The framework is intended to support reproducible, country-level hypothesis generation rather than to directly identify causal policy effects; prospective, intervention-specific studies remain necessary before using these outputs to guide resource allocation. These approaches may complement descriptive epidemiology by generating country-specific hypotheses for prospective evaluation.

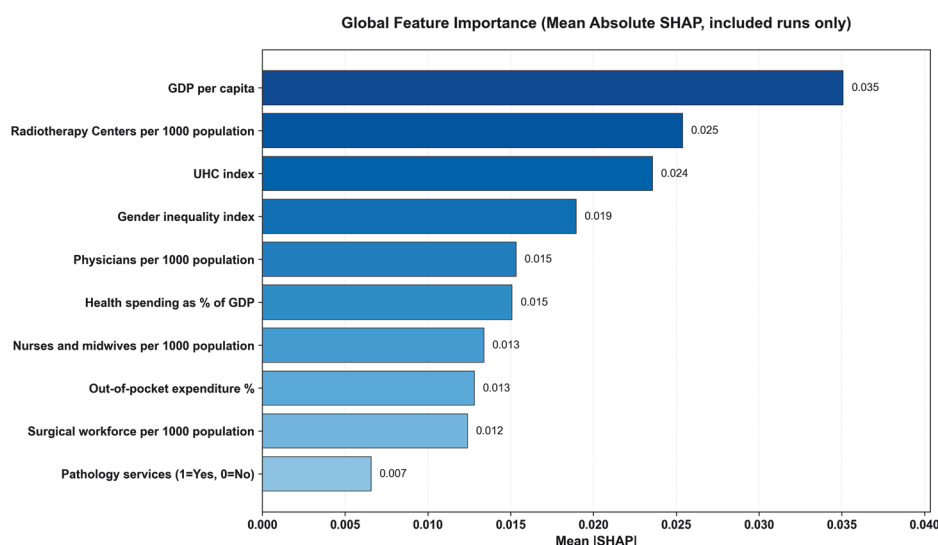


Fig. 2 Global feature importance for predicting prostate cancer MIR (mean absolute SHAP values). Bar plot showing the mean absolute SHAP value for each predictor variable, averaged across all cross-validation runs in which a given country appeared in the test fold. Mean absolute SHAP values summarize overall (global) feature importance only; per-country directional contributions for the leading predictors are illustrated in the country-level SHAP decompositions described for Yemen and the United States in the main text.

Limitations

This study has several limitations inherent to an ecological, country-level design. National MIR estimates do not reflect individual patient trajectories and cannot establish causal mechanisms. Cancer registry quality varies substantially between countries, particularly in regions with limited resources for diagnosis and patient care, and some GLOBOCAN MIR inputs are based on modeled rather than directly observed data, which may bias between-country comparisons. Although we included all 185 countries with a GLOBOCAN 2022 prostate cancer MIR estimate, this sample is still relatively small for a machine learning model and may limit predictive power and generalizability, particularly for subtle nonlinear interactions or rare covariate patterns across heterogeneous global contexts. To partially address these concerns, an income-group stratified sensitivity analysis is reported (Supplementary Figs. 1, 2). Although SHAP enhances interpretability by quantifying predictor influence, these results remain associational and hypothesis-generating, not causally prescriptive. Unmeasured confounders, including screening and PSA testing practices, oncology-specific provider density and within-country workforce distribution, health-seeking behaviors, cultural factors, and political context, as well as substantial subnational disparities in infrastructure, are not captured in this framework, and results may therefore mask important local heterogeneity. Greater data granularity, careful stratification by registry quality, integration of causal inference methods, and additional sensitivity analyses will be important in future studies to ensure robust, context-informed guidance for policy and targeted intervention. Prospective, intervention-specific evaluation remains essential before model-derived priorities can be translated into cancer control action.

CONCLUSIONS

SHAP-informed machine learning can predict country-level prostate cancer MIR with reasonable accuracy and highlights radiotherapy infrastructure and the UHC service coverage index, alongside GDP per capita, as the health system features most consistently associated with lower MIR across countries. These ecological findings suggest that targeted investment in radiotherapy capacity and in the breadth of health coverage may be a higher-priority hypothesis for prostate cancer mortality reduction

than undifferentiated increases in aggregate health spending, although this must be confirmed by intervention-specific prospective evaluation. Together, the results support continued attention to equity, access, and system-specific resource prioritization in global prostate cancer control.

REFERENCES

- Dudipala H, Jani CT, Gurnani SD, Morgenstern-Kaplan D, Tran E, Edwards K, et al. Evolving global trends in prostate cancer: disparities across income levels and geographic regions (1990–2019). *JCO Glob Oncol*. 2025;11:e2500249. <https://doi.org/10.1200/GO-25-00249>.
- Schäfer EJ, Laversanne M, Sung H, Soerjomataram I, Briganti A, Dahut W, et al. Recent patterns and trends in global prostate cancer incidence and mortality: an update. *Eur Urol*. 2025;87:302–13. <https://doi.org/10.1016/j.eururo.2024.11.013>.
- Force LM, Kocarnik JM, May ML, Bhangdia K, Crist A, Penberthy L, et al. The global, regional, and national burden of cancer, 1990–2023, with forecasts to 2050: a systematic analysis for the Global Burden of Disease Study 2023. *Lancet*. 2025;406:1565–86. [https://doi.org/10.1016/S0140-6736\(25\)01635-6](https://doi.org/10.1016/S0140-6736(25)01635-6).
- Dee EC, Wu JF, Feliciano EJG, Ting FIL, Willmann J, Ho FDV, et al. National cancer system characteristics and global pan-cancer outcomes. *JAMA Oncol*. 2025;11:650–4. <https://doi.org/10.1001/jamaoncol.2025.0473>.
- Patel MS, Pramesh CS, Sanford NN, Feliciano EJG, Nguyen PL, Iyengar P, et al. Machine learning reveals country-specific drivers of global cancer outcomes. *Ann Oncol*. 2026;37:565–73. <https://doi.org/10.1016/j.annonc.2025.11.014>.
- Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulina A. CatBoost: unbiased boosting with categorical features. 2017. <https://doi.org/10.48550/ARXIV.1706.09516>.
- Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, et al., editors. *Advances in Neural Information Processing Systems*, vol. 30, Curran Associates, Inc.; 2017.
- Ibrahim A, El Baldi M, Mohammed S, El Rhazi K, Benazzouz B. Cancer statistics in Yemen: incidence and mortality, in 2020. *BMC Public Health*. 2024;24:962. <https://doi.org/10.1186/s12889-024-18207-4>.
- Beckett M, Goethals L, Kraus RD, Denysenko K, Barone Mussalem Gentiles MF, Pynda Y, et al. Proximity to radiotherapy center, population, average income, and health insurance status as predictors of cancer mortality at the county level in the United States. *JCO Glob Oncol*. 2023;9:e2300130. <https://doi.org/10.1200/GO.23.00130>.
- Sanders J, Levell C, Brand LJ, Dahut WL, Nyame YK, Mucci L, et al. Consensus-built recommendations to improve prostate cancer outcomes: a summary of the American Cancer Society Prostate Cancer Collaborative. *Cancer*. 2025;131:e70010. <https://doi.org/10.1002/cncr.70010>.

11. Angell BJ, Prinja S, Gupta A, Jha V, Jan S. The Ayushman Bharat Pradhan Mantri Jan Arogya Yojana and the path to universal health coverage in India: overcoming the challenges of stewardship and governance. *PLoS Med.* 2019;16:e1002759 <https://doi.org/10.1371/journal.pmed.1002759>.

AUTHOR CONTRIBUTIONS

MSP conceived and designed the study, curated the data, developed the analysis code, performed the statistical analysis, generated the figures, and wrote the first draft of the manuscript. ECD supervised the study and provided senior oversight. JJ-N, FILT, JFW, SMM, DER, EJGF, HT, PI, JW, DEG, PLN, VM, BAM, NYL, HN, and ECD critically reviewed and edited the draft for important intellectual content. All authors reviewed and approved the final manuscript and agree to be accountable for all aspects of the work.

FUNDING

Dr. Dee is funded in part through the Prostate Cancer Foundation Young Investigator Award, mentored by Dr Iyengar, Dr Rathkopf, and Dr McBride. The funder had no role in the design of the study; the collection, analysis, or interpretation of data; the writing of the manuscript; or the decision to submit the manuscript for publication.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41391-026-01156-x>.

Correspondence and requests for materials should be addressed to Milit S. Patel or Edward Christopher Dee.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.