5. Fedewa SA, Gansler T, Smith R, et al. Recent Patterns in Shared Decision Making for Prostate-Specific Antigen Testing in the United States. *Annals of family medicine*. Mar 2018;16(2):139-144. doi:10.1370/afm.2200
6. Richards TB, Dai S, Gray SC, Hall IJ, Siegel DA. Number of prostate-specific antigen (PSA) screening tests in the last five years reported by men in the United States in 2010, 2015, and 2018. *Urologic oncology*. Feb 27 2022;doi:10.1016/j.urolonc.2022.01.016
7. Riviere P, Kalavacherla S, Banegas MP, et al. Patient perspectives of prostate cancer screening vary by race following 2018 guideline changes. *Cancer*. Jan 1 2023;129(1):82-88. doi:10.1002/cncr.34530
8. Kalavacherla S, Riviere P, Kalavacherla S, Anger JT, Murphy JD, Rose BS. Prostate Cancer Screening Uptake in Transgender Women. *JAMA network open*. Feb 5 2024;7(2):e2356088. doi:10.1001/jamanetworkopen.2023.56088
9. Kessler LG, Comstock B, Aiello Bowles EJ, et al. Protocol to measure validity and reliability of colorectal, breast, cervical and lung cancer screening questions from the 2021 National Health Interview Survey: Methodology and design. *PloS one*. 2024;19(3):e0297773. doi:10.1371/journal.pone.0297773
10. Guimond P, Bunn H, O'Connor AM, et al. Validation of a tool to assess health practitioners' decision support and communication skills. *Patient education and counseling*. Jul 2003;50(3):235-45. doi:10.1016/s0738-3991(03)00043-0
11. Elwyn G, Edwards A, Wensing M, Hood K, Atwell C, Grol R. Shared decision making: developing the OPTION scale for measuring patient involvement. *Qual Saf Health Care*. Apr 2003;12(2):93-9. doi:10.1136/qhc.12.2.93
12. American Urological Association. Implementation of Shared Decision Making into Urological Practice 2022. Available at https://www.auanet.org/guidelines-and-quality/quality-and-measurement/quality-improvement/clinical-consensus-statement-and-quality-improvement-issue-brief-(ccs-and-qiib)/shared-decision-making, accessed May 7, 2024.

SIP25-002 Prostate cancer active surveillance identification using electronic health records

Project Description.

A person can be categorized as a prostate cancer survivor when a prostate tissue biopsy confirms the presence of prostate cancer. Typically, low risk localized prostate cancer grows very slowly and, in many patients, may not progress to cause clinical symptoms or early mortality. Active surveillance (AS) is an ongoing regimen where the patient is closely monitored (e.g., with repeat prostate biopsy and PSA tests) to identify prostate cancer progression. Definitive treatment with intent to cure (i.e., radical prostatectomy or radiation therapy) is provided when progression is detected. Prostate cancer AS can help minimize complications or harms from unnecessary treatment with surgery or radiation therapy. Thus, men diagnosed with low risk, localized prostate cancer managed by AS can live longer, healthier lives. The American Urological Association (AUA) guidelines indicate clinicians should recommend AS as the preferred management option for prostate cancer survivors with low-risk, localized prostate cancer.[1] In the United States, community-based rates of AS have increased from 27% in 2014 to 60% in 2021, but with wide variation across practices and practitioners.[2] The use of AS is lower among African Americans and patients with lower incomes.[3-4].

For surveillance of trends, watchful waiting needs to be distinguished from prostate cancer AS. With watchful waiting, a prostate cancer case with limited life expectancy is followed with no definitive treatment and no repeat testing until the patient develops clinical symptoms. The health provider then provides palliative treatment for those symptoms with no intent to provide definitive treatment with intent to cure.[1]

It should be possible to examine prostate cancer cases managed with AS using electronic health record data. To do this, a case definition, or computable phenotype, that accurately described AS cases would need to be identified. A research challenge has been that a large amount of information in electronic health records is unstructured. The term unstructured refers to data in an electronic health record system that does not conform to a predefined data model or structure, making it difficult to categorize and analyze by computer. Common examples of unstructured information include clinical notes by providers, documents in PDF format, medical images, pathology reports, and patient correspondence. Natural Language Processing and machine learning models can be used to process and extract information from the unstructured data. Additional research is needed on how to improve and validate computable phenotypes to identify AS for prostate cancer surveillance based on the unstructured data contained in electronic health records.[5-10]

The primary purpose of this project is for the applicant to propose, develop, and evaluate improved computable phenotypes (case definitions) for use with electronic health records (EHRs) to identify whether a person with a biopsy-confirmed diagnosis of low-risk, localized prostate cancer is being managed by AS, especially phenotypes based at least in part on the unstructured data contained in electronic health records. Of special interest are valid estimates of prostate cancer AS for African American, Hispanic, non-Hispanic White, and low-income men.

Project Objectives and Outcomes

Project objectives are:

- Propose, develop, and evaluate improved computable phenotypes for use with electronic health record systems to identify whether a person with a biopsy-confirmed diagnosis of low-risk, localized prostate cancer is being managed by AS, especially phenotypes based at least in part on unstructured data contained in electronic health records. The phenotypes are expected to provide accurate and valid information about prostate cancer AS for African American, Hispanic, non-Hispanic White, and low-income men. The electronic health record system can be used at a single healthcare location or integrated across multiple sites. The applicant may propose one or several phenotypes to evaluate.
- Validate the proposed phenotypes by comparing the number and characteristics of cases (e.g., patient age, Prostate Specific Antigen level, Gleason Grade Group, clinical stage, histopathology, education, socioeconomic status, insurance coverage, and duration of management with active surveillance) identified by the proposed phenotypes with the number and characteristics of cases identified by medical chart review or by methods used by urological clinical registries (or equivalent) to extract information from electronic health records.[11-13]
- Evaluate the proposed phenotypes for any bias by the structures and processes of the health care systems providing the electronic health record data, and propose methods to mitigate bias.[14]

- Evaluate whether the proposed phenotypes can be used to predict future events such as prostate cancer progression that requires treatment with surgery or radiation.[15]

Outcomes: At the conclusion of this project, applicants are expected to provide a final report with details on the recommended phenotypes that accurately define a prostate cancer case managed with AS, a summary of any limitations (biases) of the study phenotypes, and any suggestions on how to best use Natural Language Processing and machine learning methods in future projects. In addition, applicants are expected to make available for open access (e.g., on GitHub) any final computer algorithms recommended by the project to use the recommended phenotypes or to facilitate processing of unstructured electronic health record data to identify prostate cancer active surveillance.

Healthy People 2030 Objectives
C-11: Increase the proportion of cancer survivors living 5 years or longer after diagnosis.
C-R01: Increase quality of life for cancer survivors**.**

NCCDPHP/DCPC Priority
Cancer survivors live longer, healthier lives.

Project Activities and Submission Requirements
Applications submitted in response to this SIP should present a **Research Plan** that addresses the following requirements listed below under study design and methods:

Study design and methods
The applicant will propose an appropriate study design to develop and evaluate phenotypes to identify prostate cancer cases managed with AS using data from electronic health records (EHRs), especially phenotypes based on the unstructured data electronic health records in electronic health records. The applicant will be expected to include a description of the electronic health record systems that will be analyzed, the number of cases with prostate cancer and years of data available, the structured data sources that will be included (e.g., a list of the ICD-9 and ICD-10, CPT, HCPCS, SNOMED, and LOINC codes that will be analyzed), and details on the unstructured data sources and variables that will be analyzed (e.g., clinical notes and laboratory, radiology, surgery, and laboratory PDF reports). The applicant will be expected to include details on the proposed sample size for African American, Hispanic, non-Hispanic White, and low-income men, and the methods that will be used to evaluate whether sample sizes are sufficient to answer the study questions. The applicant will be expected to include details on the methods that will be used to assess whether the proposed study phenotypes vary across education, socioeconomic status, and insurance coverage. The applicant also will be expected to include details on how Natural Language Processing and machine learning methods will be used, the characteristics of cases that will be evaluated, the methods for validation of prostate cancer active surveillance, and the methods to determine and mitigate bias by the structures and processes of the health care systems providing the electronic health record data.

Population of Focus

The study population will include male patients of any age who received their first prostate biopsy, and where the biopsy reported low-risk clinically localized prostate cancer. (The American Urological Association Guidelines define clinically localized low-risk prostate cancer as: Prostate-Specific Antigen <10 ng/mL AND Gleason Grade Group 1 AND clinical stage T1-T2a)[1]. To evaluate prostate cancer AS, men who meet the study definition will need to be followed over their lifetime from their first prostate cancer biopsy confirmed diagnosis until patient management with AS ends (e.g., the patient receives surgery or radiation therapy) or until the study observation period ends, whichever occurs first. Study populations are preferred that can provide valid estimates for Black or African American, Hispanic or Latino, non-Hispanic White, and low-income men.

Collaboration/Partnerships

Describe plans for collaboration/partnerships with data sources to develop and analyze phenotypes for active surveillance of prostate cancer and accomplish the study objectives. Examples might include state cancer registries (or equivalent), urological cancer registries (or equivalent), academic, medical practice-based sources, or other data sources. The plan should describe how partners will be engaged throughout the period of performance to inform the study.

Recruitment Plan

Applicants will be expected to obtain approvals from study health care locations to analyze and evaluate phenotypes using the electronic health record systems and data at those locations. If needed to validate the phenotypes, applicants also will be expected to obtain any additional approvals needed from patients, providers, or study locations.

Annual Action Plan

Provide a 12-month action plan using SMART goals and objectives to include a progressive timeline for completion of activities.

Evaluation Plan/Performance measurement

Applicants must provide an evaluation and performance measurement plan that demonstrates how the recipient will fulfill the requirements described. The evaluation plan must meet SMART goals and be consistent with the CDC evaluation framework (https://www.cdc.gov/evaluation/). A plan to evaluate data gathered as part of the research plan should be included.

Data Management Plan

If the applicant is collecting public health data, a standalone data management plan that addresses the 5 elements of AR-25 must be submitted in the Other Plan(s) section of the application.
https://www.cdc.gov/grants/additional-requirements/ar-25.html Applicants should use the NCCDPHP Data Management Plan template available at:
https://www.cdc.gov/nccdphp/dch/media/files/Data-Management-Plan-template.docx

Dissemination & Translation Plan

Provide details on anticipated strategies to disseminate and translate the findings of the research; for example, any anticipated peer-reviewed scientific articles, conferences, or other plans to

distribute project highlights to public health practitioners, non-governmental organizations, and healthcare decision makers. Applicants are expected to make available for open access (e.g., on GitHub) any final computer algorithms recommended by the project to use the recommended phenotypes or to facilitate processing of unstructured electronic health record data to identify prostate cancer AS.

Public Health Impact
The computable phenotypes for electronic health records potentially may lead to improved data to plan and make interventions to increase prostate cancer AS and to reduce disparities in the use of prostate cancer AS.[16-17] In addition, the project findings on methods to extract information from unstructured electronic health record data may be useful as part of other projects to modernize public health data and provide actionable insights for decision-making at all levels of public health and healthcare.

Special Eligibility and Responsiveness
The following criteria specific to this SIP will be used to determine the institution's eligibility and responsiveness:

Special Eligibility Requirements:

Access to the numbers of the subpopulations of interest needed to address the study questions.

Responsiveness Criteria:

The applicant must provide documentation (e.g., letters of support or memorandum of agreement) demonstrating evidence of access to the numbers of the subpopulations of interest needed to address the study questions.

Additional Review Criteria
In addition to the standard review criteria (Significance, Approach, Innovation, Investigators, and Environment) used to evaluate the scientific and technical merit of research applications, the following additional review criteria specific to this SIP will be considered in the determination of scientific merit and the priority score:

- Do the investigators show previous research expertise that have provided high quality outputs and contributed to improvements in public health practice?
- Does the applicant provide evidence of expertise of team members on AS of prostate cancer, clinical phenotypes based on electronic medical records, including Natural Language Processing, machine learning, and statistical modeling?
- Does the applicant describe experience creating translation and dissemination products for public health practitioners and non-governmental organizations?
- Does the proposed study design include details on the study sample size, especially for African American, Hispanic, non-Hispanic White, and low-income men?
- Does the applicant include details on the proposed methods to assess whether the results of the proposed phenotypes vary across education, socioeconomic status, and insurance coverage?

Funding Preferences

The following preferences specific to this SIP will be considered in the funding decision:

None

Research Plan Length and Supporting Material

The Research Strategy Section of the Research Plan is limited to a maximum of 12 pages. Supporting material included as appendices may not exceed 10 PDF (maximum of 30 pages) attachments. The applicant must provide documentation, e.g., letters of support or memorandum of agreement, demonstrating evidence of access or assuring that the applicant has access to the electronic health record data sets that the applicant has proposed to analyze to develop and evaluate phenotypes.

Availability of Funds

It is anticipated that approximately **$800,000** is available to fund 1 Prevention Research Center for a 2-**year** period of performance. The average award for each recipient is expected to be approximately **$400,000** for year one. The year one ceiling per recipient is **$400,000**. Funding may vary and is subject to change. Funding available includes direct and indirect costs.

Research Status

This project will be non-exempt research involving human subjects. It is anticipated that this project will require local IRB approval. Applicants should provide a federal- wide assurance number for each performance site.

OMB/PRA

OMB/PRA is not expected to apply

Award Administration

CDC staff will serve as consultants on this project, and will provide technical assistance, as requested, on project activities such as evaluation design, data collection and analysis, and data interpretation and dissemination of results. CDC staff may be co-authors on manuscripts. However, CDC staff will not have contact with human subjects and will not receive or analyze identifiable data.

References

1. Eastham JA, Auffenberg GB, Barocas DA, et al. Clinically localized prostate cancer: AUA/ASTRO Guideline, Part I: introduction, risk assessment, staging, and risk-based management. J Urol. 2022;208(1):10-18. doi: 10.1097/JU.0000000000002757. PMID: 35536144.

2. Cooperberg MR, Meeks W, Fang R, et al. Time trends and variation in the use of active surveillance for management of low-risk prostate cancer in the US. JAMA Netw Open. 20231;6(3):e231439. doi: 10.1001/jamanetworkopen.2023.1439. PMID: 36862409.

3. Al Hussein Al Awamlh B, Wu X, et al. Intensity of observation with active surveillance or watchful waiting in men with prostate cancer in the United States. Prostate Cancer Prostatic Dis. 2023;26(2):395-402. doi: 10.1038/s41391-022-00580-z. PMID: 35882950.

4. Walton EL, Deebajah M, Keeley J, et al. Barriers to obtaining prostate multi-parametric magnetic resonance imaging in African American men on active surveillance for prostate cancer. Cancer Med. 2019;8(8):3659-3665. doi: 10.1002/cam4.2149. PMID: 31111654
5. Modi PK, Kaufman SR, Qi J, et al. National trends in active surveillance for prostate cancer: validation of Medicare claims-based algorithms. Urology. 2018;120:96-102. doi: 10.1016/j.urology.2018.06.037. PMID: 29990573
6. Knighton AJ, Belnap T, Brunisholz K, et al. Using electronic health record data to identify prostate cancer patients that may qualify for active surveillance. EGEMS (Wash DC). 2016;4(3):1220. doi: 10.13063/2327-9214.1220. PMID: 27683663.
7. Hernandez-Boussard T, Kourdis PD, Seto T, et al. Mining electronic health records to extract patient-centered outcomes following prostate cancer treatment. American Medical Informatics Assoc. Annual Symposium Proc. 2018;2017:876-882. PMID: 29854154.
8. Boyd RA, Hilton CA, Duke JD. A Gleason score extractor for electronic health records. Poster. 2019 American Medical Informatics Association (AMIA) Clinical Informatics Conference. Atlanta, GA. 2019.
9. Yu S, Le A, Feld E, Schriver E, et al. A Natural Language Processing-assisted extraction system for Gleason scores: development and usability study. JMIR Cancer. 2021 Jul 2;7(3):e27970. doi: 10.2196/27970. PMID: 34255641.
10. Zhang Y, Cai T, Yu S, et al. High-throughput phenotyping with electronic medical record data using a common semi-supervised approach (PheCAP). Nat Protoc. 2019;14(12):3426-3444. doi: 10.1038/s41596-019-0227-6. PMID: 31748751
11. Michigan Urological Surgery Improvement Collaborative (MUSIC) Registry. https://musicurology.com/
12. American Urological Association (AUA) Quality (AQUA) Registry https://www.auanet.org/research-and-data/aua-quality-(aqua)-registry
13. Canary Prostate Active Surveillance Study (PASS) Registry. https://canarypass.org/
14. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization. 2021. https://iris.who.int/bitstream/handle/10665/341996/9789240029200-eng.pdf
15. Newcomb LF, Schenk JM, Zheng Y, et al. Long-term outcomes in patients using protocol-directed active surveillance for prostate cancer. JAMA. 2024 May 30:e246695. doi: 10.1001/jama.2024.6695. PMID: 38814624.
16. Gaylis F, Cohen E, Calabrese R, Prime H, et al. Active surveillance of prostate cancer in a community practice: how to measure, manage, and improve? Urology. 2016;93:60-7. doi: 10.1016/j.urology.2015.11.053. PMID: 26876460.
17. Carlsson SV, Preston M, Vickers A, et al. Provider perceptions of an electronic health record prostate cancer screening tool. Appl Clin Inform. 2024;15(2):282-294. doi: 10.1055/s-0044-1782619. PMID: 38599619.

SIP25-003 Scaling What Works within the National Comprehensive Cancer Control Program

Project Description

CDC's National Comprehensive Cancer Control Program (NCCCP) has a 26-year history of working to enhance large scale efforts to reduce cancer risk, improve screening utilization, enhance