Cervical Cancer Screening Research in the PROSPR I Consortium: Rationale, Methods, and Baseline Findings from a U.S. Cohort

Aruna Kamineni[1], Jasmin A. Tiro[2,3], Elisabeth F. Beaber[4], Michael J. Silverberg[5], Cosette M. Wheeler[6], Chun R. Chao[7], Jessica Chubak[1], Celette Sugg Skinner[2,3], Douglas A. Corley[5], Jane J. Kim[8], Bijal A. Balasubramanian[3,9], V. Paul Doria-Rose[10] PROSPR consortium

[1]Kaiser Permanente Washington Health Research Institute, Seattle, WA

[2]University of Texas Southwestern Medical Center, Dallas, TX

[3]Simmons Comprehensive Cancer Center, Dallas, TX

[4]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA

[5]Division of Research, Kaiser Permanente Northern California, Oakland, CA

[6]University of New Mexico Comprehensive Cancer Center, Albuquerque, NM

[7]Department of Research and Evaluation, Kaiser Permanente Southern California, Pasadena, CA

[8]Center for Health Decision Science, Harvard T.H. Chan School of Public Health, Boston, MA

[9]UTHealth School of Public Health in Dallas, Dallas, TX

[10]Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD

Abstract

Little is known about the effect of evolving risk-based cervical cancer screening and management guidelines on United States (US) clinical practice and patient outcomes. We describe the National Cancer Institute's Population-based Research Optimizing Screening through Personalized Regimens (PROSPR I) consortium, methods, and baseline findings from its cervical sites: Kaiser Permanente Washington, Kaiser Permanente Northern California, Kaiser Permanente Southern California, Parkland Health & Hospital System/University of Texas Southwestern (Parkland-UTSW), and New Mexico HPV Pap Registry housed by University of New Mexico (UNM-NMHPVPR). Across these diverse healthcare settings, we collected data on human papillomavirus (HPV) vaccinations, screening tests/results, diagnostic and treatment procedures/results, and cancer diagnoses on nearly 4.7 million women aged 18–89 years from 2010–2014. We calculated baseline (2012 for UNM-NMHPVPR; 2010 for other sites) frequencies for sociodemographics, cervical cancer risk factors, and key screening process measures for each site's cohort. Healthcare delivery settings, cervical cancer screening strategy, race/ethnicity, and insurance status varied among sites. The proportion of women receiving a Pap test during the baseline year was similar

---

Corresponding author: Aruna Kamineni, PhD, MPH, Kaiser Permanente Washington Health Research Institute, Metropolitan Park East, 1730 Minor Ave. Ste 1600, Seattle, WA 98101, Fax: 206-287-2485 Aruna.S.Kamineni@kp.org.

across sites (26.1–36.1%). Most high-risk HPV tests were performed either reflexively or as co-tests, and utilization pattern varied by site. Prevalence of colposcopy or biopsy was higher at Parkland-UTSW (3.6%) than other sites (1.3–1.4%). Incident cervical cancer was rare. HPV vaccination among age-eligible women not already immunized was modest across sites (0.1–7.2%). Cervical PROSPR I makes available high-quality, multilevel, longitudinal screening process data from a large and diverse cohort of women to evaluate and improve the effectiveness of US cervical cancer screening delivery.

Keywords

cervical cancer; human papillomavirus; cancer screening; Pap test

Introduction

The dramatic reduction of cervical cancer morbidity and mortality through screening with the Papanicolaou (Pap) test followed by diagnosis and treatment of precursor lesions and early cancers (collectively the cervical cancer screening process[1]) is one of the greatest achievements for women's health care in the past 75 years.[2–4] Cervical cancer prevention has evolved as knowledge about the natural history of this cancer has grown and the etiologic role of human papillomavirus (HPV) has been incorporated into primary and secondary prevention strategies.[5–7] Tests for high-risk HPV (hrHPV) types and the first HPV vaccine became available in 2003 and 2006, respectively. Integrating both updated knowledge and these newer technologies, United States (US) guideline organizations now recommend cervical cancer screening between the ages of 21–65 for average-risk women, with different screening intervals based on age and test modality.[5–7] These recommendations are a significant departure from the long-established practice of annual Pap testing beginning at initiation of sexual intercourse for all women, and represent a shift toward screening for cervical cancer based on individual risk (i.e., risk-based screening). Advances in cervical cancer prevention technologies and changes in screening guidelines have also been accompanied by shifts in management guidelines for abnormal results,[5,8,9] which currently consider a host of individual risk factors including age and prior test results.

With the rapid evolution of new technologies and changing screening and management guidelines, we have limited knowledge about their collective effect on US clinical practice and ultimately, patient outcomes. We need data to: 1) understand patient-, provider-, facility-, and system-level variation in delivering screening and managing abnormal results; 2) identify challenges and determine solutions to implementation of risk-based screening and management algorithms; and 3) guide quality improvement efforts. These data, particularly when linked to patient outcomes, can inform groups developing clinical practice guidelines who seek to optimize the balance of benefits and harms in their recommendations as well as health care systems working to effectively implement risk-based strategies into clinical practice. Population-based Research Optimizing Screening through Personalized Regimens (PROSPR I), a National Cancer Institute initiative, provides a rich data resource to address these key knowledge gaps in a country lacking a national health care system,

universal health care coverage, centrally-organized screening, and central registries for data collection and monitoring activities.

Initiated in 2011, the PROSPR I consortium includes a Statistical Coordinating Center and breast, cervical, and colorectal Research Centers.[1,10–12] The four cervical Research Centers, comprising five sites, represent a variety of US health care settings. They provide a unique opportunity to examine the effect of heterogeneous clinical practices on cervical cancer prevention during a period of evolving technologies and changing guidelines. This manuscript describes: 1) the overall design, objectives, setting, and cohort for cervical PROSPR I; 2) the methods used by cervical Research Centers to collect and harmonize data on key elements of the cervical cancer screening process; 3) the comprehensive quality assurance, documentation, and management activities critical to the development, utilization, and stewardship of PROSPR I's central data repository; and 4) baseline findings from the cervical PROSPR I cohort. This description is intended to provide researchers interested in collaborating with PROSPR I investigators with information on available resources and facilitate comparisons with other, similar data.

Materials and Methods

Cervical PROSPR I Design, Objectives, Setting, and Cohort

Three PROSPR I Research Centers—Kaiser Permanente Washington (KPWA, formerly Group Health), Kaiser Permanente Northern California (KPNC) and Kaiser Permanente Southern California (KPSC) combined (KPNC/KPSC), and Parkland Health & Hospital System/University of Texas Southwestern (Parkland-UTSW)—are both PROSPR colorectal[12] and cervical Research Centers; the New Mexico HPV Pap Registry (NMHPVPR) housed by University of New Mexico (UNM-NMHPVPR), is a cervical Research Center only. The primary objectives of the cervical Research Centers were to: 1) collect high-quality, multilevel, longitudinal cervical cancer screening process data on defined populations; 2) submit common data elements (CDEs) to create a central data repository; and 3) collaborate on cervical-specific and trans-organ research that elucidates patient-, provider-, facility-, and system-level factors influencing the screening process. The Statistical Coordinating Center, located at Fred Hutchinson Cancer Research Center, led data harmonization activities, quality assurance and management of the central data repository, statistical analyses, and communication across the consortium. Institutional Review Boards across sites and the Statistical Coordinating Center approved all research activities.

Located in the Southern and Western regions of the US, the cervical PROSPR I Research Centers have diverse health care settings and patient populations. KPWA is a mixed-model, nonprofit health plan serving nearly 651,000 Washington State residents. The KPNC/KPSC Research Center comprises two integrated health care systems delivering care to approximately 8 million members in Northern and Southern California. Parkland-UTSW is the sole safety-net provider for underinsured and uninsured Dallas County, Texas residents, delivering care to approximately 167,000 adults annually. The UNM-NMHPVPR Research Center is a unique statewide, population-based cervical cancer screening registry.[13] Under state mandate, results from all Pap and HPV tests as well as diagnostic and treatment pathology from New Mexico residents are reported to the registry.[14] In these health care

settings, most employer-based/private insurance plans cover HPV vaccination and cervical cancer screening without patient cost-sharing as mandated by the Affordable Care Act of 2010.[15] Publicly-financed programs also cover cervical cancer screening services (i.e., initial screening test, diagnostic evaluation, and treatment) and include: Medicare, a national insurance program for those over age 65; Medicaid, a joint federal and state insurance program for low income individuals under age 65 (eligibility varies by state); and federal block grants for uninsured women who do not qualify for Medicaid (e.g., National Breast and Cervical Cancer Early Detection Program, Title X family planning program).[16] In terms of HPV vaccination, the federal Vaccines for Children Program[17] fully covers vaccination for children ages 18 and under who are Medicaid-eligible, uninsured, American Indian or Alaska Native, or underinsured. Depending on the state, uninsured adults may have vaccine coverage through Medicaid or the Section 317 Immunization Grant Program,[18] a federal block grant.

Cervical PROSPR I cohort definitions varied slightly across sites. KPWA, KPNC, KPSC, and Parkland-UTSW included women beginning on January 1, 2010, while the UNM-NMHPVPR cohort began on January 1, 2012. All cohorts include women aged 18–89 years except Parkland-UTSW, which truncates at age 64 because residents who become Medicare-eligible at age 65 may opt to receive care outside the Parkland-UTSW system. To ensure capture of test results, KPWA included women in the Group Practice Division of the health plan (i.e., members who selected or were assigned to a KPWA Medical Center for their primary care). Parkland-UTSW included women with at least one primary care or women's health visit. As a registry site, UNM-NMHPVPR included only women who had undergone cervical cancer screening, diagnosis, or treatment. During 2010–2014, the combined cervical cohort included nearly 4.7 million women.

Data Collection

The central data repository containing longitudinal breast, colorectal, and cervical screening process data is the core resource of PROSPR I, enabling pooled analyses of data harmonized across sites. The repository contains data from approximately 6.6 million patients and includes 148 cervical-specific and 124 trans-organ CDEs[19,20] encompassing characteristics at the patient, provider, clinic, facility, and system or registry level. The cervical Research Centers collected information on demographics, pregnancies, cervical cancer risk factors, the cervical cancer screening process (screening tests, diagnostic evaluations, treatment procedures, and outcomes), cervical cancer screening history prior to cohort entry, cancer diagnoses, and characteristics of providers and facilities. Data collection largely relied on extraction from clinical and administrative data sources, and for UNM-NMHPVPR, from the registry information systems. The electronic health record (EHR) is the primary clinical data source at KPWA, KPNC, KPSC, and Parkland-UTSW. EHRs make patient health information readily accessible at the point of care to authorized providers, enabling clinical decision-making and care coordination. EHRs contain patient health data recorded by providers and may include self-reported information from patient questionnaires. This documentation by providers serves as a legal record of health care, and facilitates billing and reimbursement. Below we describe the methods used by Research Centers to collect and harmonize data on key cervical cancer screening process elements (Table 1).

HPV vaccinations.—KPWA, KPNC and KPSC maintain internal databases that include both provider-administered and patient-reported immunizations. In addition to these sources, KPWA also has records identified from utilization data from external providers and the Washington State Immunization Information System.[21] Parkland-UTSW collects both provider-administered and patient-reported immunizations from the EHR. These sites electronically extracted HPV vaccine administration dates, valency, and number of doses on cohort members from their data sources. Vaccine data from UNM-NMHPVPR were unavailable.

Screening tests.—KPWA, KPNC, KPSC, and Parkland-UTSW identified dates of Pap and HPV tests from the EHR, with some Parkland-UTSW records from a legacy clinical database. UNM-NMHPVPR collected this information from the registry. All sites electronically extracted Pap results, including specimen adequacy. Results were mapped to Bethesda[22] CDE categories via electronic text analysis at KPSC and UNM-NMHPVPR. KPWA, KPNC, and Parkland-UTSW electronically mapped semi-structured results and manually reviewed a limited set of records that were complex (i.e., less structured) or non-standardized (e.g., containing non-Bethesda terms). All sites recorded the most severe result when a Pap test had multiple results. Extraction and mapping of structured HPV results were achieved electronically at KPWA, KPNC, and KPSC. Parkland-UTSW performed manual review to link HPV results from the legacy database to EHR information. UNM-NMHPVPR performed routine periodic manual review for quality assurance (QA) given the large number of laboratories serving New Mexico's population and the wide variety of HPV tests implemented in recent years.

Diagnostic evaluations and treatment procedures.—To harmonize data collection, KPWA, KPNC, KPSC, and Parkland-UTSW collaborated to create a comprehensive list of Healthcare Common Procedure Coding System (HCPCS), Current Procedural Terminology (CPT-4), and International Classification of Diseases-Ninth Revision (ICD-9) codes for cervical diagnosis and treatment procedures. Codes identify colposcopies, biopsies (including endocervical curettage and endocervical brushing), ablative procedures (cryotherapy and laser therapy), excisional treatments (loop electrosurgical excision procedure and conization), and gynecological surgeries (trachelectomy/cervicectomy, hysterectomy, and pelvic exenteration) and when applicable, were supplemented with health care system-specific codes. Using this comprehensive code list, KPWA, KPNC, and KPSC extracted dates for diagnostic and treatment procedures from utilization data and Parkland-UTSW extracted identical information from clinical orders and visit data. At KPWA and Parkland-UTSW, procedure type was confirmed or modified using information from pathology reports linked to biopsy and excisional procedures by unique medical record numbers and specimen collection dates. Pathology reports are electronically reported to the NMHPVPR. From this source, UNM-NMHPVPR extracted dates for all diagnosis and treatment procedures except ablations and colposcopies without biopsies or excisions, which are not available in the registry. Pathology reports from biopsies, excisions, and surgeries were available from the NMHPVPR information systems or from the EHR at the other sites. Across Research Centers, unstructured histopathology diagnoses were mapped to clinically meaningful categories, with the most severe diagnosis recorded for each procedure.

Classification was by manual review of all pathology reports at KPWA, by natural language parsing algorithms at UNM-NMHPVPR,[23,24] and by electronic text analysis methods at KPNC, KPSC, and Parkland-UTSW. For validation, Parkland-UTSW manually reviewed at least 10% of each category; KPNC and KPSC performed several iterative reviews that included all cancer diagnoses and up to 100 records from other result categories; UNM-NMHPVPR reviewed all cancer diagnoses, and a randomly selected sample from other result categories.

Cervical cancer outcomes.—All sites linked to high-quality cancer registries to identify invasive cervical cancer diagnoses, tumor characteristics (e.g., stage and histology), and first course of cancer treatment among cohort members. KPWA linked to the Seattle-Puget Sound Surveillance, Epidemiology, and End Results[25] (SEER) registry using a matching algorithm with EHR number, social security number, birthdate, sex, and last name. UNM-NMHPVPR linked to the New Mexico SEER registry using Registry Plus™ Link Plus,[26] a probabilistic record linkage program developed by the US Centers for Disease Control and Prevention. KPNC and KPSC linked to their SEER-contributing cancer registries using unique EHR numbers. The KPNC and KPSC cancer registries follow SEER practices in case ascertainment and characterization of histopathology, invasiveness, tumor size, extension, and lymph node involvement. Parkland-UTSW provided first and last names, sex, and address history to the Texas Cancer Registry,[27] a member of the North American Association of Central Cancer Registries, for probabilistic linkage using Registry Plus™ Link Plus[26] and SAS software.

Data Harmonization, Quality Assurance, and Documentation

The Statistical Coordinating Center's data harmonization approach has been described previously.[28] Data QA was integrated into all processes and continually reviewed for improvement. In collaboration with the Research Centers, the Statistical Coordinating Center developed a request packet for data that included CDE specifications, recommended quality control steps for the Research Centers, and data submission instructions. Upon receiving requested data from the Research Centers, the Statistical Coordinating Center processed the data and generated a QA report with CDE distributions, probability plots, extreme values, non-permissible values, and logic checks. The Statistical Coordinating Center and Research Centers reviewed reports and identified any additional data issues. Research Centers submitted revised data if necessary, QA reports were regenerated, and the process continued until all data issues were resolved. Following this process, data were transformed to harmonized datasets containing derived variables that accounted for differences in data structures across Research Centers (e.g., actual dates vs. time since a reference date). If additional data issues were identified during analyses, revised data were submitted or other CDEs were added, thereby supplementing the primary data QA process. The Statistical Coordinating Center also evaluated data comparability across and within Research Centers over time to ensure credible, high-quality harmonized PROSPR I data.

Data documentation was critical for both data QA and management given the heterogeneity in data sources and collection methods across Research Centers. The Statistical Coordinating Center collected and organized documentation at the level of CDEs, data files,

and Research Centers. This data documentation included cohort definitions, characteristics of Research Center data sources, programming code used for mapping when available, any deviations from CDE definitions or considerations for interpretation, and source data dictionaries or forms, if applicable. Additionally, the Statistical Coordinating Center documented each Research Center's data sources and mapping rules used to extract data and create CDEs.

Analysis

We calculated baseline (2012 for UNM-NMHPVPR and 2010 for the other Research Centers) frequencies for sociodemographic characteristics, cervical cancer risk factors, and select screening process measures for each site's cohort. For the registry site (UNM-NMHPVPR), US Census data were used to estimate the population, age distribution, race/ethnicity counts, and Rural-Urban Commuting Area[29,30] measures. To report vaccine uptake during the baseline year for the KPWA, KPNC, KPSC, and Parkland-UTSW cohorts, we calculated the proportion of women who received one or more doses among those who had zero doses at cohort entry and were vaccine age-eligible through the end of 2010. To characterize cytology and HPV testing, colposcopic evaluations, and incident cervical cancer diagnoses during the baseline year, we calculated the proportion of women who received each procedure or a cancer diagnosis among those enrolled in 2012 for UNM-NMHPVPR and 2010 for the other Research Centers. For these measures, we identified the at-risk populations at KPWA, KPNC, KPSC, and Parkland-UTSW by excluding women with EHR documentation of an absent cervix or history of invasive cervical cancer. We used the publicly-available 2012 Behavioral Risk Factor Surveillance System (BRFSS)[31]survey-reported estimate of 21.9% hysterectomy prevalence among 18- to 84-year-old New Mexico women and calculated the at-risk population for the UNM-NMHPVPR cohort by multiplying the census population estimate by $1 - 0.219$, but were not able to exclude women with a history of invasive cervical cancer. To report incident cervical cancers in the KPWA cohort, we additionally restricted the denominator to women residing in a Seattle-Puget Sound SEER county for the entire baseline year.

Results

Clinical Environment of the Cervical PROSPR I Research Centers

The five sites comprising the cervical Research Centers are heterogeneous in geographic region, type of health care delivery setting, and strategy for delivering cervical cancer screening (Table 2). Four sites are integrated health care delivery systems or have a predominantly integrated model with primary care practices affiliated with specialty providers. This model enabled systematic tracking of the entire cervical cancer screening process at these sites. UNM-NMHPVPR captures cervical cytopathology and histopathology data from all health care delivery settings in New Mexico. All five sites began offering either the bivalent or quadrivalent HPV vaccine by 2007, with the nonavalent vaccine available by 2016. Co-testing (concomitant performance of Pap and HPV testing) was adopted as the primary screening strategy at KPNC in 2003 and KPSC in 2005. Co-testing has been available as an option at KPWA since 2012, and at Parkland-UTSW and in New Mexico since 2004. All sites currently use liquid-based cytology, although time of adoption ranged

from 2000 to 2011, and the processing method differs across centers (KPWA, KPNC, and KPSC use SurePath™, Parkland-UTSW uses ThinPrep®, and both liquid-based cytology methods as well as conventional cytology are used in New Mexico). Methods for HPV testing also vary across sites. Screening is delivered opportunistically at Parkland-UTSW and largely so in New Mexico, while KPWA, KPNC, and KPSC remind women to get screened through organized programs. KPWA, KPNC, KPSC, and Parkland-UTSW employ an Epic®-based EHR with programmed electronic alerts to notify providers during a clinic visit if a patient is due or overdue for cervical cancer screening.

Characteristics of the PROSPR I Cervical Cohort by Research Center

Baseline sociodemographic characteristics, cervical cancer risk factors, and key screening process measures for cohort members from each of the cervical Research Centers are described in Table 3.

Sociodemographic characteristics.—Age distributions of cohort members at KPWA, KPNC, KPSC, and UNM-NMHPVPR are similar, although UNM-NMHPVPR has a slightly larger proportion of 66- to 89-year-olds. However, even after accounting for truncation at age 64 (i.e., by restricting to those 65 years across all sites), the Parkland-UTSW cohort is younger than other sites with nearly 60% under 40 years (data not shown). Race and ethnicity vary substantially across Research Centers. KPWA enrollees are predominantly non-Hispanic white (75.0%), with a lower proportion at KPNC (54.3%), although they are still the largest race/ethnic group. In contrast, most of Parkland-UTSW patients are Hispanic (62.6%) and one-quarter are non-Hispanic black. The largest racial and ethnic groups at KPSC and UNM-NMHPVPR comprise non-Hispanic whites and Hispanics, with equal distributions of the two groups within each cohort. KPNC has the highest proportion of Asians and Pacific Islanders (18.5%) and UNM-NMHPVPR has the highest proportion of Native Americans or Alaskan Natives (8.3%).

All or almost all cohort members across the integrated health care systems live in metropolitan areas. In contrast, 34.6% of women in New Mexico live in lower-density areas. KPWA, KPNC, and KPSC are very similar in insurance coverage distribution—most patients are commercially insured (76.9–83.0%) or have Medicare (14.2–20.5%), and very small proportions have Medicaid or are dual-eligible for Medicare and Medicaid. In contrast, many of Parkland-UTSW patients (41.4%) are uninsured (medical care costs are covered by Dallas County's medical assistance program financed by property taxes) and over 30% have cervical cancer screening coverage through federal block grants (National Breast and Cervical Cancer Early Detection Program,[32] family planning programs).[33] Nearly 20% of Parkland-UTSW patients have Medicaid coverage which, in Texas, is mostly restricted to pregnant women. Individual-level insurance data were unavailable for UNM-NMHPVPR; however, estimates using US Census data suggest that, in 2013, a year most closely aligned with UNM-NMHPVPR's baseline, approximately 23% of New Mexico women aged 19–64 were uninsured or had Indian Health Service coverage only and 18% were covered by Medicaid.[34]

Cervical cancer risk factors.—Similar proportions of women aged 18–89 had EHR documentation of surgery to remove the cervix at KPWA, KPNC, and KPSC (8.5–10.0%). The proportion was lower at Parkland-UTSW (5.2%). Individual-level information was not available for the UNM-NMHPVPR cohort, but the proportion with prior hysterectomy among 18- to 84-year-old New Mexico women was estimated at 21.9% (range: 20.5–23.2%) using self-reported data from the 2012 BRFSS survey.[31] The prevalence of a prior cervical cancer diagnosis at the integrated health care systems was very low (0.1–0.3%). The prevalence of human immunodeficiency virus (HIV) infection at the integrated health care systems was also very low, with the largest proportion at Parkland-UTSW (1.2%). Vaccinated women were a modest proportion of the integrated health care system cohorts at baseline (0.3–6.9%). However, after restricting to women who were age-eligible to be vaccinated during the vaccine era (18–29 years old at cohort entry), the proportions vaccinated were 30.0% at KPWA, 19.4% at KPNC, 29.2% at KPSC, and 1.1% at Parkland-UTSW (data not shown).

Vaccinations, Pap and HPV testing, diagnostic evaluations, and cervical cancer outcomes in the baseline year.—At KPWA, KPNC, KPSC, and Parkland-UTSW, HPV vaccine uptake during 2010 among age-eligible women (18–26 years through the end of 2010) who had not already been vaccinated prior to cohort entry was modest, ranging from 0.1–7.2%. Among the integrated health care systems, the proportion of women in the cohort (aged 18–64 at Parkland-UTSW and aged 18–89 at KPWA, KPNC, and KPSC) with no EHR documentation of an absent cervix and no history of cervical cancer who received a Pap test ranged from 26.1–36.1% during 2010. After accounting for the publicly-available 2012 BRFSS survey-based estimate of hysterectomy prevalence,[31] we estimated that 30.4% of at-risk women aged 18–89 in New Mexico received a Pap test in 2012. Approximately 21% of these women received a hrHPV test at KPNC and KPSC, but much smaller proportions of women at KPWA (0.9%), Parkland-UTSW (6.4%), and in New Mexico (6.3%) were HPV-tested during the baseline year. Nearly 83% of hrHPV tests were performed reflexively at KPWA. In contrast, most or nearly all were part of co-tests at the other Research Centers. Similar proportions of 18- to 89-year-old women underwent either a colposcopic examination at KPWA, KPNC, and KPSC or a colposcopy yielding a biopsy in New Mexico (1.3–1.4%), with a higher prevalence of colposcopy at Parkland-UTSW (3.6%). Incident cervical cancer diagnoses in the baseline year were rare, with the largest proportion observed at Parkland-UTSW. Age-stratified estimates of the proportions of the cohorts receiving cytology tests, hrHPV tests, and colposcopic evaluations during the baseline year are provided in a supplemental table.

Discussion

Representing demographically, geographically, and organizationally diverse US health care settings, cervical PROSPR I offers a rich data resource to evaluate and improve US cervical cancer screening delivery. The depth and breadth of the comprehensive cervical cancer screening process data collected from PROSPR I Research Centers can enable investigation of key research priorities in cervical cancer screening including: 1) sources of variation and points in the process vulnerable to failure across health care settings; and 2) adherence to

evolving and complex screening and management guidelines. In addition, PROSPR I data enables exploration of how newer technologies and management algorithms improve effectiveness of the screening process, and benefits and harms of screening strategies by risk status. In the cervical PROSPR I cohort, the large numbers of racial and ethnic minority groups (Asians=409,304; blacks=291,060; Hispanics=1,136,083), and those who are uninsured (n=34,721) or receive Medicaid coverage (n=96,145) enables investigation of populations that may experience screening process disparities. Furthermore, unlike other large, national datasets such as the National Health Interview Survey[35] or the Behavioral Risk Factor Surveillance System[36] that rely on self-reported information from participants, PROSPR relies on more objective information from health plan administrative data, EHRs and supporting clinical information systems, and registry information systems from the Research Centers. Additionally, all eligible patients are included in PROSPR, not just those who volunteer to participate.

The cervical PROSPR I data have some limitations. Characterizing the cervical cancer screening process across multiple sites with heterogeneous clinical information systems via secondary data collection is inherently challenging, complicating the analysis and interpretation of pooled data. The implementation of EHRs, combined with advances in informatics such as increased computational power and methods such as natural language processing to capture unstructured data offer great potential for research, especially for large, national consortia such as PROSPR. However, leveraging data collected or recorded for varied purposes requires knowledge of the original intended use and evaluation of whether research use is appropriate.[37] A multitude of factors influence the specific content of data produced by EHRs across settings and over time, including: 1) local clinical workflows; 2) legal requirements and reimbursement incentives (e.g., Meaningful Use and Healthcare Effectiveness Data and Information Set [HEDIS]); and 3) the EHR vendor's data model and local implementation. Thus, data provenance—the context in which clinical data are recorded—complicates use of EHR data for research. The cervical Research Centers worked to account for provenance during data collection to ensure appropriate capture and interpretation of screening process data.

Data availability varies across health care settings and local data structures and sources may also constrain mapping to CDEs. As part of data QA and management, the cervical Research Centers and the Statistical Coordinating Center worked closely together to minimize these constraints and document them when unavoidable. Data completeness may depend on membership history, health care utilization, and documentation of care received outside of the health care system. For example, HPV vaccination prevalence in the cervical PROSPR I cohorts may not reflect true prevalence since immunizations received outside of the health care systems may be less complete for new enrollees, those with limited utilization history, or in states (e.g., Texas) with opt-in immunization registries.[38,39] Published age-specific hysterectomy prevalence estimates are not available for New Mexico and this is a limitation of the results presented in the supplemental table. Additionally, hysterectomy prevalence as ascertained by EHR documentation of an absent cervix in the KPWA, KPNC, and KPSC cohorts appears to be lower than some national survey estimates.[40] Despite substantial efforts by these health care systems to ascertain this information to improve accuracy of HEDIS reporting and ensure that only at-risk women are included in screening reminder

programs, EHR documentation of surgical history prior to health plan enrollment may be incomplete. However, given that documentation was available for very similar proportions of women across the three systems, and may be more reliable than self-reported hysterectomy status, the data may in fact be consistent with the true prevalence in these cohorts. Information on socioeconomic indicators such as education or income is not routinely collected for either health plan administration or clinical care, and individual-level data are not available from the US Census. Age at initiation of sexual intercourse and lifetime number of sexual partners are also rarely ascertained at visits in these health care settings. Although smoking status is ascertained during care delivery, these data are not currently available in the PROSPR I central data repository and is a limitation of this resource.

The administrative and clinical environment of the health care setting may provide additional context to explain similarities and differences in cancer screening process measures across the cervical PROSPR I sites. For example, lower uptake of HPV vaccination in two sites during 2010 may reflect reduced outreach for immunizations among 18- to 26-year-olds at KPNC, and is consistent with the lack of public payor programs to reimburse vaccinations in young adults at Parkland-UTSW. The higher prevalence of HPV testing in the baseline year at KPNC and KPSC reflects their earlier and systematic adoption of co-testing. Far fewer HPV tests than Pap tests were done at Parkland-UTSW in 2010, and most were performed for surveillance in women who had had a prior abnormal cytology result. The higher proportions of colposcopic evaluations and cervical cancer diagnoses at Parkland-UTSW likely reflects a combination of factors including younger age of the cohort, higher HPV-related abnormalities, and disparities in access to preventive care (i.e., collectively resulting in higher cervical cancer risk). These examples illustrate the importance of considering variation in clinical implementation, administrative constraints, and system-level policies when monitoring trends in the cervical cancer screening process within a health care system and pooling data across systems. The cervical Research Centers took great care to ascertain and document system-level factors contributing to variation as an adjunct to quantitative data collection to facilitate interpreting data and appropriately contextualizing results.

During the PROSPR I study period, US guidelines for both cervical cancer screening and management of abnormal screening results were modified to incorporate risk factors such as age, modality-specific test results, and prior screening and histology results.[5–9] While these newer risk-based guidelines are intended to maximize benefits and minimize harms[41] associated with screening and follow-up, acceptance by both women and providers of the more complex recommendations is not well understood, particularly those calling for longer screening intervals or managing patients with conflicting Pap and HPV test results. It is also unclear whether implementation of the modified guidelines has improved cervical outcomes since delivering guideline-concordant care and monitoring delivery is now much more complex. Adding further complexity to this landscape is that, unlike other high-resource countries with a single-payor system, health care in the US is complex and fragmented into a multitude of payors (e.g., government, employer, self), payment models (e.g., fee-for-service/procedure-based, fee-for-quality/disease-based), and organizations delivering health care services (e.g., health maintenance organizations [HMOs], accountable care organizations [ACOs]. Since most US cervical cancer screening delivery occurs in the

absence of the call-recall programs available in countries with government-sponsored health care, monitoring screening coverage and management of abnormal results will be essential as we strive to optimize cervical cancer screening, diagnosis, and treatment.

PROSPR I provides a rare opportunity to understand and advance cervical cancer prevention and control during a period of remarkable change in which the US system of health care is confronted by new prevention technologies and strategies, guideline consensus, and payment reform. Ultimately, the benefits of cervical cancer screening can be realized only if it is delivered well. The screening process data available from the PROSPR I consortium provide collaborative opportunities to evaluate increasingly risk-based screening and management of abnormal results through the lens of clinical implementation. The diversity of the pooled cervical PROSPR I cohort in demographics, health care settings, and insurance coverage will strengthen the external validity of study findings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Abbreviations:

BRFSS	Behavioral Risk Factor Surveillance System
CDEs	common data elements
CPT-4	Current Procedural Terminology
EHR	electronic health record
HCPCS	Healthcare Common Procedure Coding System
HEDIS	Healthcare Effectiveness Data and Information Set
HPV	human papillomavirus
hrHPV	high-risk HPV
ICD-9	International Classification of Diseases-Ninth Revision
KPNC	Kaiser Permanente Northern California
KPSC	Kaiser Permanente Southern California
KPWA	Kaiser Permanente Washington
NMHPVPR	New Mexico HPV Pap Registry

Pap	Papanicolaou test
Parkland-UTSW	Parkland Health & Hospital System/University of Texas Southwestern
PROSPR I	Population-based Research Optimizing Screening through Personalized Regimens
QA	quality assurance
SEER	Surveillance, Epidemiology, and End Results
UNM-NMHPVPR	University of New Mexico-New Mexico HPV Pap Registry
US	United States

References

1. Beaber EF, Kim JJ, Schapira MM, Tosteson AN, Zauber AG, Geiger AM, Kamineni A, Weaver DL, Tiro JA, Population-based Research Optimizing Screening through Personalized Regimens Consortium. Unifying screening processes within the PROSPR consortium: a conceptual model for breast, cervical, and colorectal cancer screening. J Natl Cancer Inst 2015;107(6):djv120. [PubMed: 25957378]

2. Landis S, Murray T, Bolden S, Wingo P. Cancer statistics, 1999. CA Cancer J Clin 1999;49(1):8–31, 31. [PubMed: 10200775]

3. Wingo PA, Cardinez CJ, Landis SH, Greenlee RT, Ries LA, Anderson RN, Thun MJ. Long-term trends in cancer mortality in the United States, 1930–1998. Cancer 2003;97(12 Suppl):3133–3275. [PubMed: 12784323]

4. International Agency for Research on Cancer, World Health Organization. Cervix cancer screening. Vol 10 Lyon, France: IARC Press, 2005.

5. Committee on Practice Bulletins-Gynecology. ACOG Practice Bulletin Number 131: screening for cervical cancer. Obstet Gynecol 2012;120(5):1222–1238. [PubMed: 23090560]

6. Moyer VA US Preventive Services Task Force. Screening for cervical cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2012;156(12):880–891, W312. [PubMed: 22711081]

7. Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, Garcia FA, Moriarty AT, Waxman AG, Wilbur DC, Wentzensen N, Downs LS Jr., Spitzer M, Moscicki AB, Franco EL, Stoler MH, Schiffman M, Castle PE, Myers ER, ACS-ASCCP-ASCP Cervical Cancer Guideline Committee. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. CA Cancer J Clin 2012;62(3):147–172. [PubMed: 22422631]

8. ACOG Committee on Practice Bulletins--Gynecology. ACOG practice bulletin no. 109: Cervical cytology screening. Obstet Gynecol 2009;114(6):1409–1420. [PubMed: 20134296]

9. Massad LS, Einstein MH, Huh WK, Katki HA, Kinney WK, Schiffman M, Solomon D, Wentzensen N, Lawson HW, 2012 ASCCP Consensus Guidelines Conference. 2012 updated consensus guidelines for the management of abnormal cervical cancer screening tests and cancer precursors. J Low Genit Tract Dis 2013;17(5 Suppl 1):S1–S27. [PubMed: 23519301]

10. National Cancer Institute, Division of Cancer Control & Population Sciences. Population-based Research Optimizing Screening through Personalized Regimens (PROSPR) home page. 2018; http://healthcaredelivery.cancer.gov/prospr/. Accessed July 30, 2018.

11. Onega T, Beaber EF, Sprague BL, Barlow WE, Haas JS, Tosteson AN, D Schnall M, Armstrong K, Schapira MM, Geller B, Weaver DL, Conant EF. Breast cancer screening in an era of personalized regimens: a conceptual model and National Cancer Institute initiative for risk-based and

preference-based approaches at a population level. Cancer 2014;120(19):2955–2964. [PubMed: 24830599]

12. Tiro JA, Kamineni A, Levin TR, Zheng Y, Schottinger JS, Rutter CM, Corley DA, Skinner CS, Chubak J, Doubeni CA, Halm EA, Gupta S, Wernli KJ, Klabunde C. The colorectal cancer screening process in community settings: A conceptual model for the population-based research optimizing screening through personalized regimens consortium. Cancer Epidemiol Biomarkers Prev 2014;23(7):1147–1158. [PubMed: 24917182]

13. Center for HPV Prevention. About the New Mexico HPV pap registry. https://hpvprevention.unm.edu/NMHPVPR/. Accessed July 30, 2018.

14. New Mexico Department of Health. Notifiable diseases or conditions in New Mexico. 2016; http://nmhealth.org/publication/view/regulation/372/. Accessed July 30, 2018.

15. Strokoff S, Grossman E, Sterkx C, Blount E, Volberding E, Office of the Legislative Counsel for the use of the U.S. House of Representatives. Compilation of Patient Protection and Affordable Care Act [as amended through May 1, 2010] including Patient Protection and Affordable Health Care Act health-related portions of the Health Care and Education Reconciliation Act of 2018. 111th Congress 2d Session. 2010; http://housedocs.house.gov/energycommerce/ppacacon.pdf. Accessed July 31, 2018.

16. Corley DA, Haas JS, Kobrin S. Reducing variation in the "standard of care" for cancer screening: recommendations from the PROSPR Consortium. JAMA 2016;315(19):2067–2068. [PubMed: 27089237]

17. Centers for Disease Control and Prevention. Vaccines for Children Program (VFC). 2018; https://www.cdc.gov/vaccines/programs/vfc/index.html. Accessed July 30, 2018.

18. Centers for Disease Control and Prevention. Questions answered on vaccines purchased with 317 funds. For immunization Managers. 2013; https://www.cdc.gov/vaccines/imz-managers/guides-pubs/qa-317-funds.html. Accessed July 30, 2018.

19. Nadkarni PM, Brandt CA. The common data elements for cancer research: remarks on functions and structure. Methods Inf Med 2006;45(6):594–601. [PubMed: 17149500]

20. CTEP common data elements. National Cancer Institute, NCI Wiki. 2016; https://wiki.nci.nih.gov/display/caDSR/CTEP+Common+Data+Elements. Accessed April 27, 2017.

21. Washington State Health Department of Health. Information for healthcare providers about the immunization information system. https://www.doh.wa.gov/ForPublicHealthandHealthcareProviders/HealthcareProfessionsandFacilities/DataReportingandRetrieval/ImmunizationInformationSystem/ForProviders. Accessed July 30, 2018.

22. Foreword Solomon D.. In: Nayar R, Wilbur DC, eds. The Bethesda System for Reporting Cervical Cytology: Definitions, Criteria, and Explanatory Notes. 3rd ed New York: Springer, 2015:v–vii.

23. Gage J, Schiffman M, Hunt W, Joste N, Ghosh A, Wentzensen N, Wheeler C, New Mexico HPV Pap Registry Steering Committee. Cervical histopathology variability among laboratories: a population-based statewide investigation. Am J Clin Pathol 2013;139(3):330–335. [PubMed: 23429369]

24. Benard V, Castle P, Jenison S, Hunt W, Kim J, Cuzick J, Lee J, Du R, Robertson M, Norville S, Wheeler C, New Mexico HPV Pap Registry Steering Committee. Population-based incidence rates of cervical intraepithelial neoplasia in the human papillomavirus vaccine era. JAMA Oncol 2017;3(6):833–837. [PubMed: 27685805]

25. National Cancer Institute. Surveillance, epidemiology, and end results program data, 1973–2012 http://seer.cancer.gov/data/. Accessed July 30, 2018.

26. Division of Cancer Prevention and Control, Centers for Disease Control and Prevention. Link Plus. National Program of Cancer Registries. 2015; https://www.cdc.gov/cancer/npcr/tools/registryplus/lp_tech_info.htm. Accessed July 30, 2018.

27. Texas Department of State Health Services. Data linkages. Texas Cancer Registry. 2018; http://www.dshs.texas.gov/tcr/data/linkages.aspx?terms=linkage. Accessed July 30, 2018.

28. Rolland B, Reid S, Stelling D, Warnick G, Thornquist M, Feng Z, Potter JD. Toward rigorous data harmonization in cancer epidemiology research: one approach. Am J Epidemiol 2015;182(12):1033–1038. [PubMed: 26589709]

29. WWAMI Rural Health Research Center. Rural-urban commuting area codes. http://depts.washington.edu/uwruca/index.php. Accessed July 30, 2018.

30. United States Department of Agriculture, Economic Research Service. Rural-urban commuting area codes. 2016; http://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes.aspx. Accessed February 8, 2018.

31. Centers for Disease Control and Prevention. BRFSS 2012 survey data and documentation 2014; http://www.cdc.gov/brfss/annual_data/annual_2012.html. Accessed July 30, 2018.

32. Centers for Disease Control and Prevention. National Breast and Cervical Cancer Early Detection Program. 2018; https://www.cdc.gov/cancer/nbccedp/about.htm. Accessed July 30, 2018.

33. Nikpay SS. Federal support for family planning clinics associated with dramatic gains in cervical cancer screening. Womens Health Issues 2016;26(2):176–182. [PubMed: 26781749]

34. State health facts: health insurance coverage of women 19–64. 2015; http://kff.org/other/state-indicator/nonelderly-adult-women/. Accessed July 8, 2018.

35. Centers for Disease Control and Prevention. National Health Interview Survey. National Center for Health Statistics. 2018; https://www.cdc.gov/nchs/nhis/index.htm. Accessed July 30, 2018.

36. Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System. 2018; https://www.cdc.gov/brfss/index.html. Accessed July 30, 2018.

37. Johnson KE, Kamineni A, Fuller S, Olmstead D, Wernli KJ. How the provenance of electronic health record data matters for research: a case example using system mapping. EGEMS (Washington, DC) 2014;2(1):1058.

38. Boom JA, Sahni LC, Nelson CS, Dragsbaek AC, Franzini L. Immunization information system opt-in consent: at what cost? J Public Health Manag Pract 2010;16(5):E18–25.

39. Collins BK, Morrow HE, Ramirez JM, Cochran CE, Smith DR. Childhood immunization coverage in US states: the impact of state policy interventions and programmatic support. J Health Soc Policy 2006;22(1):77–92. [PubMed: 17135110]

40. Rositch AF, Nowak RG, Gravitt PE. Increased age and race-specific incidence of cervical cancer after correction for hysterectomy prevalence in the United States from 2000 to 2009. Cancer 2014;120(13):2032–2038. [PubMed: 24821088]

41. Marcus PM, Pashayan N, Church TR, Doria-Rose VP, Gould MK, Hubbard RA, Marrone M, Miglioretti DL, Pharoah PD, Pinsky PF, Rendle KA, Robbins HA, Roberts MC, Rolland B, Schiffman M, Tiro JA, Zauber AG, Winn DM, Khoury MJ. Population-based precision cancer screening: a symposium on evidence, epidemiology, and next steps. Cancer Epidemiol Biomarkers Prev 2016;25(11):1449–1455. [PubMed: 27507769]

Novelty and Impact:

Population-based Research Optimizing Screening through Personalized Regimens (PROSPR I), an initiative of the United States National Cancer Institute, established a data repository for cervical cancer screening research. Data include human papillomavirus vaccinations, screening tests/results, diagnostic and treatment procedures/results, and cancer diagnoses on nearly 4.7 million women aged 18–89 years from 2010–2014 in Washington, California, New Mexico, and Texas. This article describes setting, data collection, and baseline characteristics for the PROSPR I cervical cohort.

Table 1.

Data Sources and Collection Methods for Key Screening Process Data Elements by Cervical PROSPR I Research Center

Screening Process Data Elements	Kaiser Permanente Washington	Kaiser Permanente		Parkland Health & Hospital System-University of Texas Southwestern	University of New Mexico – New Mexico HPV Pap Registry
		Northern California	Southern California		
HPV Vaccinations [*date, valency, and dose*]	Electronic extraction from EHR of KPWA-administered, patient-reported, utilization, and Washington State Immunization Information System data	Electronic extraction from vaccination database and EHR of KP-administered and patient-reported data		Electronic extraction from EHR of Parkland-administered and patient-reported data	---
Tests [*date*]					
Cytology	Electronic extraction from EHR (laboratory/pathology databases)			Electronic extraction from EHR and Clinical Data Repository[a]	Electronic extraction from NMHPVPR information systems
HPV					
Diagnostic Evaluations [*date*]					
Colposcopy	Electronic extraction from utilization data			Electronic extraction from clinical orders with confirmation from utilization data and pathology reports, when available	---
Biopsy	Electronic extraction from utilization data with confirmation from manual review of linked pathology reports	Electronic extraction from utilization data			Electronic extraction from NMHPVPR information systems
Treatments [*date*]					
Cryotherapy	Electronic extraction from utilization data			Electronic extraction from EHR orders with confirmation from utilization data and pathology reports	---
Laser therapy					
Conization	Electronic extraction from utilization data with confirmation from manual review of linked pathology reports	Electronic extraction from utilization data			Electronic extraction from NMHPVPR information systems
LEEP					
Surgery					
Results from Tests and Procedures					
Cytology	Electronic text-based extraction from EHR aided by manual review of complex or non-standard text strings	Electronic text-based extraction from laboratory database aided by manual review to guide approach for non-standard text strings and quality assurance	Text analysis of cytology reports from pathology database	Electronic text-based extraction from EHR aided by manual review	Natural language text parsing of cytology reports from NMHPVPR information systems.
HPV	Electronic extraction from EHR	Electronic extraction from laboratory databases			Electronic extraction from NMHPVPR information systems with routine periodic manual review for quality assurance

Screening Process Data Elements	Kaiser Permanente Washington	Kaiser Permanente		Parkland Health & Hospital System- University of Texas Southwestern	University of New Mexico – New Mexico HPV Pap Registry
		Northern California	Southern California		
Biopsy/ECC/ECB	Manual review of pathology reports from pathology database and EHR	Electronic text-based extraction of pathology reports from laboratory/pathology databases with sampling for verification by manual review			Natural language text parsing from NMHPVPR information systems with periodic targeted manual review for quality assurance
Treatment histology					
Cervical cancer diagnoses [*date, tumor type, and stage*]	Linkage to SEER	Linkage to SEER-contributing KPNC Cancer Registry	Linkage to SEER-contributing KPSC Cancer Registry	Linkage to Texas Cancer Registry	Linkage to SEER

Abbreviations: EHR, electronic health record; ECB, endocervical brushing; ECC, endocervical curettage; HPV, human papillomavirus; KPWA, Kaiser Permanente Washington; KPNC, Kaiser Permanente Northern California; KPSC, Kaiser Permanente Southern California; LEEP, loop electrosurgical excision procedure; NMHPVPR, New Mexico HPV Pap Registry; Pap, Papanicolaou; SEER, Surveillance, Epidemiology, and End Results program of the National Cancer Institute.

*a*Legacy database storing some cytology and HPV tests performed prior to 2011.

Table 2.

Clinical Characteristics of Cervical PROSPR I Research Centers

		Kaiser Permanente Washington	Kaiser Permanente Northern California & Kaiser Permanente Southern California		Parkland Health & Hospital System-University of Texas Southwestern	University of New Mexico – New Mexico HPV Pap Registry
Geographic region		Washington State	Northern California	Southern California	Dallas County, Texas	State of New Mexico
System type		Mixed-model health care system with a large integrated group practice	Integrated health care system		Integrated safety-net health care system	Registry
Screening-eligible target population		Average-risk women aged 21 to 65 years				Women aged 21 to 65 years
	HPV vaccination	Quadrivalent vaccine covered since 2006 and nonavalent vaccine since May 2015	Quadrivalent vaccine covered since 2006 and nonavalent vaccine since August 2015		Quadrivalent vaccine covered since 2007 and nonavalent vaccine since June 2016	Bivalent or quadrivalent vaccine available since 2007 and nonavalent vaccine available beginning May 2015
	Conventional cytology	Not in use since switch to liquid-based cytology			Used in women's health clinics until Fall 2011	Used by some clinics
	Liquid-based cytology	SurePath™ processing method adopted in 2006	SurePath™ processing method adopted in 2009	SurePath™ processing method adopted in 2011	ThinPrep® processing method adopted in primary care clinics in 2004 and women's health clinics by late 2011	ThinPrep® or SurePath™ methods available beginning in 2000
Primary and secondary prevention strategies and screening technologies	HPV assay	Digene HC2™ introduced in 2006, Cervista™ in late 2009, and Roche cobas® in mid-2012	Digene HC2™ introduced in 2001	Digene HC2™ adopted in 2005	Digene HC2™ introduced in 2004, Cervista™ adopted in late 2010	Multiple tests used with Digene HC2™ predominating since 2003 and Roche cobas® use increasing beginning in 2012 and Hologic Aptima® use increasing beginning in 2013
	Reflex HPV testing	Introduced in 2006	Provided during 2001–2012 for 25- to 29-year-olds	Adopted in 2005 for 21- to 29-year-olds and provided only for 25- to 29-year-olds since 2013	Introduced in primary care clinics in 2004 and women's health centers in September 2011	Introduced in 2003
	HPV co-testing	Available as an option in 2012 for 30- to 65-year-olds	Recommended as primary strategy in 2003 for 30- to 65-year-olds, and beginning in 2013, for 25- to 65-year-olds	Adopted as primary strategy in 2005 for 30- to 65-year-olds	Introduced as an option in 2004 for 30- to 65-year-olds, policies encourage use for symptomatic patients and no Pap test history in primary care clinics and surveillance in women's health centers	Available as an option in 2004

	Kaiser Permanente Washington	Kaiser Permanente Northern California & Kaiser Permanente Southern California		Parkland Health & Hospital System-University of Texas Southwestern	University of New Mexico – New Mexico HPV Pap Registry
Screening intervals	Cytology: 3 years Co-testing: 5 years	Cytology: 3 years Co-testing: 3 years	Cytology: 3 years Co-testing: 5 years	Cytology: 3 years Co-testing: 5 years	Variable
Screening outreach/inreach	Eligible women receive reminders via a "birthday letter." Overdue women receive telephone outreach from their primary care clinic, and the EHR contains automated alerts to support opportunistic care, persisting until a screening test is ordered or the alert is overridden.	Eligible women receive electronic or mailed reminders through a regional tracking system that identifies women due for screening. Overdue women receive additional electronic or emailed reminders from their providers, and the EHR contains automated alerts to support opportunistic care.	Eligible women receive mailed reminders. Overdue women receive automated telephone calls from the health plan, and the EHR contains automated alerts to support opportunistic care.	Eligible women are offered screening during clinic visits. The EHR has an automated alert to remind providers to ascertain screening status and offer screening if due/overdue.	Variable by practice, facility, or organization.
Providers offering screening		Primary care and Ob/Gyn[a]			
EHR system		EPIC			Variable
Pathology system	Progeny	Cerner CoPath Plus		Cerner Millenium	Variable
Cancer registry	Cancer Surveillance System[b,c]	Kaiser Permanente Northern California Cancer Registry	Kaiser Permanente Southern California Cancer Registry	Texas Cancer Registry[c]	New Mexico Tumor Registry[b,c]

Abbreviations: EHR, electronic health record; HC2, Hybrid Capture 2; HPV, human papillomavirus; Pap, Papanicolaou.

[a]At Parkland Health & Hospital System, infectious disease specialists provide primary care services including cervical screening for most patients with human immunodeficiency virus.

[b]Surveillance, Epidemiology, and End Results Program of the National Cancer Institute.

[c]North American Association of Central Cancer Registries member.

Table 3.

Demographic Characteristics, Risk Factors, and Select Screening Process Measures at Baseline in the Cervical PROSPR I Research Centers

	Kaiser Permanente Washington	Kaiser Permanente Northern California	Southern California	Parkland Health & Hospital System- University of Texas Southwestern	University of New Mexico – New Mexico HPV Pap Registry
Cohort Definition	18 to 89-year-old females enrolled in the Group Practice Division of the health plan at any time during January 1, 2010 through December 31, 2014	18 to 89-year-old females enrolled in the health plan at any time during January 1, 2010 through December 31, 2013		18 to 64-year-old female Dallas County residents with a primary care clinic or women's health center visit at any time during January 1, 2010 through September 30, 2014	18 to 89-year-old females residing in New Mexico who underwent cervical cancer screening, diagnosis, or treatment any time during January 1, 2012 through June 30, 2014
Total population in baseline year[a]	207,999	1,402,797	1,468,883	84,253	800,220
Demographics and Risk Factors	n (%)	n (%)	n (%)	n (%)	n (%)
Age in years[a]					
18–20	12,982 (6.2)	89,685 (6.4)	107,507 (7.3)	4,455 (5.3)	43,020 (5.4)
21–29	31,076 (14.9)	206,143 (14.7)	233,488 (15.9)	21,363 (25.4)	126,667 (15.8)
30–39	33,227 (16.0)	248,536 (17.7)	263,112 (17.9)	23,446 (27.8)	126,446 (15.8)
40–49	37,003 (17.8)	261,071 (18.6)	278,065 (18.9)	16,045 (19.0)	129,365 (16.2)
50–59	43,630 (21.0)	259,170 (18.5)	269,701 (18.4)	13,348 (15.8)	148,006 (18.5)
60–65[b]	21,149 (10.2)	126,285 (9.0)	125,313 (8.5)	5,596 (6.6)	78,674 (9.8)
66–89	28,932 (13.9)	211,907 (15.1)	191,697 (13.1)	---	148,042 (18.5)
Race/ethnicity[a]					
Non-Hispanic white	127,322 (75.0)	699,530 (54.3)	510,010 (38.9)	7,560 (9.0)	354,942 (44.4)
Non-Hispanic black	7,821 (4.6)	107,178 (8.3)	142,590 (10.9)	21,171 (25.2)	12,300 (1.5)
Hispanic	9,522 (5.6)	236,735 (18.4)	492,979 (37.6)	52,594 (62.6)	344,253 (43.0)
Asian/Pacific Islander	17,192 (10.1)	237,785 (18.5)	138,436 (10.5)	2,491 (3.0)	13,400 (1.7)
Native American/ Alaskan Native	1,427 (0.8)	4,701 (0.4)	2,428 (0.2)	97 (0.1)	66,549 (8.3)
Other/Multiracial	6,529 (3.8)	2,197 (0.2)	26,270 (2.0)	74 (0.1)	8,776 (1.1)
Unknown	38,186	114,671	156,170	266	0
Rural-Urban Continuum measure[c]					
Metropolitan	199,644 (96.6)	1,338,896 (95.9)	1,447,722 (98.7)	84,253 (100)	523,608 (65.4)
Micropolitan	4,993 (2.4)	38,734 (2.8)	15,900 (1.1)	0 (0.0)	196,465 (24.6)
Low density	2,050 (1.0)	18,975 (1.4)	3,513 (0.2)	0 (0.0)	80,147 (10.0)
Unknown	1,312	6,192	1,748	0	0
Health Insurance[d, e]					
Medicaid	5,942 (2.9)	32,160 (2.3)	42,037 (2.9)	16,006 (19.1)	---
Medicare	35,873 (17.2)	287,470 (20.5)	208,285 (14.2)	2,907 (3.5)	---

	Kaiser Permanente Washington	Kaiser Permanente Northern California	Kaiser Permanente Southern California	Parkland Health & Hospital System- University of Texas Southwestern	University of New Mexico – New Mexico HPV Pap Registry
Commercial/Private	159,969 (76.9)	1,082,391 (77.2)	1,218,420 (83.0)	4,401 (5.3)	---
Other Government	6,215 (3.0)	0 (0.0)	0 (0.0)	25,770 (30.7)	---
Uninsured/Charity	0 (0.0)	0 (0.0)	0 (0.0)	34,721 (41.4)	---
Unknown	0	776	141	448	---
Absent cervix documented in EHR[d, f]	18,123 (8.7)	119,252 (8.5)	147,516 (10.0)	4,351 (5.2)	---
Prior cervical cancer diagnosis[d]	421 (0.2)	1,150 (0.1)	994 (0.1)	218 (0.3)	---
HIV positive[d]	221 (0.1)	715 (0.1)	632 (0.0)	1,044 (1.2)	---
Prior HPV vaccinations[d]					
None	194,681 (93.6)	1,344,929 (95.9)	1,367,805 (93.1)	83,959 (99.7)	---
1 or more doses	13,318 (6.4)	57,868 (4.1)	101,078 (6.9)	294 (0.3)	---
Screening Process Measures in Baseline Year			2010		2012
	n (%)	n (%)	n (%)	n (%)	n (%)
Vaccinated[d, g]	1,318 (7.0)	3,242 (2.2)	10,837 (7.2)	16 (0.1)	---
Cytology tested[h]	49,525 (26.1)	338,479 (26.1)	363,760 (27.1)	28,822 (36.1)	189,910 (30.4)
hrHPV tested[h, i]	1,773 (0.9)	274,932 (21.2)	288,540 (21.5)	5,094 (6.4)	39,508 (6.3)
Reflex	1,329 (82.8)	11,902 (4.3)	5,584 (2.1)	635 (12.5)	9,386 (23.3)
Co-test	215 (13.4)	263,030 (95.7)	268,513 (97.9)	4,173 (81.8)	27,393 (67.9)
Other	61 (3.8)	0 (0.0)	0 (0.0)	286 (5.7)	2,050 (5.1)
Unknown	168	0	14,443	0	679
Evaluated with colposcopy[h, j]	2,579 (1.4)	16,781 (1.3)	17,402 (1.3)	2,854 (3.6)	8,999 (1.4)
Incident cervical cancer diagnosis[h, k]	15 (0.009)	85 (0.007)	111 (0.008)	23 (0.029)	66 (0.011)

Abbreviations: EHR, electronic health record; HPV, human papillomavirus; hrHPV, high-risk HPV; UNM-NMHPVPR, University of New Mexico-New Mexico HPV Pap Registry.

[a] Data source for UNM-NMHPVPR is United States Census data. All census records coded as Hispanic are aggregated in the Hispanic column. All other race groups represent counts for non-Hispanics. This coding of Hispanic deviates from race/ethnicity coding previously reported using New Mexico census data but is consistent with race/ethnicity construction for other PROSPR I Research Centers. Due to limitations in how age is presented in census data, race/ethnicity counts are reported for all women from New Mexico over the age of 18 as estimated for 2012.

[b] Age group for Parkland Health & Hospital System-University of Texas Southwestern (Parkland-UTSW) cohort is 60- to 64-year-olds because of cohort eligibility criteria.

[c] Rural-Urban Commuting Area (RUCA) frequencies for UNM-NMHPVPR were constructed from census 5-year population estimates by age, gender, and ZIP Code Tabulation Areas for the period ending in 2012. Percentages from census 5-year population estimates were applied to the 2012 census population totals reported for age and race/ethnicity to produce RUCA population estimates.

[d] Information was not available for UNM-NMHPVPR.

[e]Individuals in multiple insurance categories were categorized according to the following hierarchy: Medicaid, Medicare, commercial/private, other government (block grants that support cervical cancer screening at Parkland Health & Hospital System and the Basic Health Program in Washington State), uninsured/charity, and unknown.

[f]Includes women who had EHR documentation of the following surgeries prior to cohort entry: pelvic exenteration, radical hysterectomy, total hysterectomy, trachelectomy/cervicectomy, and hysterectomy, not otherwise specified.

[g]Eligible population (denominator) includes women age-eligible to receive the vaccine (18–26 years) who were not vaccinated prior to cohort entry.

[h]At-risk populations (denominators) for Kaiser Permanente Washington (KPWA), Northern California, and Southern California; and Parkland-UTSW identified by excluding women with an absent cervix or history of invasive cervical cancer. At-risk population at UNM-NMHPVPR obtained by multiplying the estimate of the percentage of women with intact cervix (1 minus the percentage of women with absent cervix [0.219] as estimated from the 2012 Behavioral Risk Factor Surveillance System) with the census population estimate.

[i]HPV test indication frequencies were calculated among those who were hrHPV tested. HPV test indication for UNM-NMHPVPR was assigned using an algorithmic approach: HPV tests 28 days following a Pap test were considered reflex or co-tests. HPV tests >28 days following a Pap test were considered to have other indications. HPV tests with unknown indication were without any preceding Pap test.

[j]Colposcopies for UNM-NMHPVPR are restricted to those with biopsies or excisions.

[k]Denominator for KPWA is also restricted to women residing in a Seattle-Puget Sound Surveillance, Epidemiology, and End Results county for the entire baseline year.