



# SURVEILLANCE RESEARCH PROGRAM

e-Newsletter  
Summer 2017



## Welcome

### SRP Associate Director's Message

Greetings from the Surveillance Research Program! We've been pretty busy since our last newsletter. Our most notable change has been our new branch, the Data Analytics Branch, as well as many new staff members. This issue includes the latest organizational chart, introductions on new staff, and several other updates and highlights from our program.

Included in our usual sections, "Initiatives" and "New on the Net," I'd like to direct your attention to a few significant announcements. Under "Initiatives," you can read all about Pilot 3 of our collaboration with the Department of Energy. This partnership will impact the future of both cancer research and scientific computing. Under "New on the Net," we highlight our brand new blog and updates to several of our resources: Cancer Trends Progress Report, SEER data, videos, software and tools. Since our

last newsletter, we have collaborated with CDC, ACS, and NAACCR on the Annual Report to the Nation on the Status of Cancer (ARN); this year's special section was on survival rates by race/ethnicity, state, and stage. The SEER website hosted the ARN microsite, which includes shareable and tweetable graphics and statistics.

We are looking forward to seeing many of you at our SEER\*DMS Meeting and at NAACCR this summer! As always, thanks for your collaboration.

Sincerely,

Lynne Penberthy  
Associate Director  
Surveillance Research Program

<https://surveillance.cancer.gov/> | <https://seer.cancer.gov/>

## In This Issue

### Updates

- New SRP Organizational Structure
- Job Opportunities
- New Staff

### Initiatives

- NCI-DOE Collaboration

### New on the 'Net

- SRP Blog
- Cancer Trends Progress Report
- Annual Report to the Nation
- SEER\*Stat
- NCI Map Stories
- SEER\*Explorer
- Did You Know? Video Series
- Twitter

### Data

- New Data Release

### Events

- NCI NLP Workshop
- SEER\*DMS Face-to-Face Meeting

### Publications and Grants

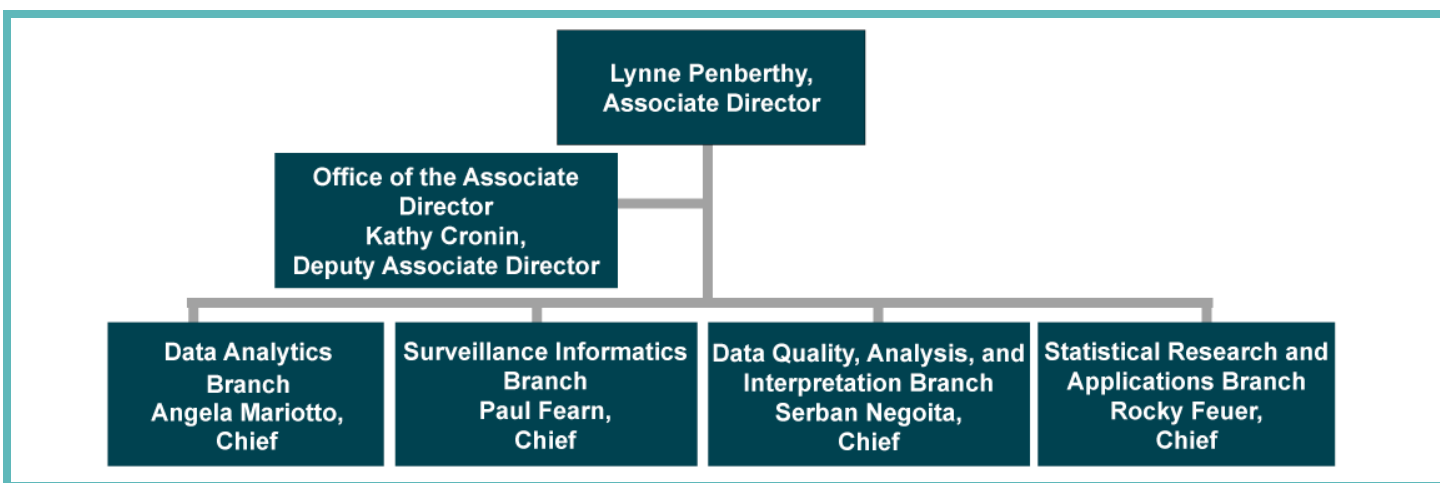
- Publications
- Grants

# UPDATES

## • New Organizational Structure

### SRP has a new branch—the Data Analytics Branch (DAB)!

We now have four branches, including the Data Analytics Branch (DAB), Surveillance Informatics Branch (SIB), Data Quality, Analysis, and Interpretation Branch (DQAIB), and Statistical Research and Applications Branch (SRAB). This reorganization reflects our major functions and directions for the future of cancer surveillance. Our new structure and corresponding mission statements are now available on [surveillance.cancer.gov](http://surveillance.cancer.gov).



### Data Analytics Branch

Our mission is to advance the science on analytic, modeling, reporting and interpretation methods for cancer epidemiological surveillance data in order to measure progress in reducing the US cancer burden.

## • Job Opportunities

### Applications are being accepted for several positions in SRP:

- CRTA Fellow for User-Driven Design and Evaluation for Cancer Surveillance
- CRTA Fellow for Project Management

For full descriptions of each position and application information, visit [surveillance.cancer.gov/jobs/](http://surveillance.cancer.gov/jobs/).



Photo Credit: NIH

# MEET the STAFF

## UPDATES: NEW STAFF

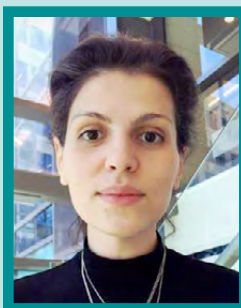
**PAUL FEARN**  
PhD, MBA



**Paul Fearn** has joined SRP as the Chief of the Surveillance Informatics Branch. He works with the SEER registries, SRP team, and collaborators to advance the application of natural language processing, machine learning, and other informatics tools and methods to support cancer surveillance.

Previously, he was Director of Biomedical Informatics at Fred Hutchinson Cancer Research Center and a founding leader of the Hutch Integrated Data Repository and Archive (HIDRA). Paul has a BA in Spanish from the University of Houston, biostatistics training from the University of Texas School of Public Health in Houston, an MBA from the New York University Stern School of Business, and a PhD in Biomedical and Health Informatics from the University of Washington School of Medicine. He has more than 20 years of experience in cancer research informatics at Baylor College of Medicine, MSKCC, Fred Hutch, and with the NCI SEER program.

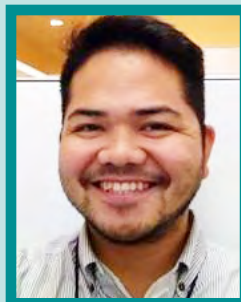
**MARINA MATATOVA**  
MA



**Marina Matatova** has joined the Surveillance Informatics Branch of SRP as the Technical Project Manager for the NCI-DOE collaboration. She works with the SRP team and SRP partners to build a scalable infrastructure that can incorporate machine-learned natural language processing tools to support precision cancer surveillance.

Marina grew up in Brooklyn, New York. She received both her Bachelor of Arts and Master of Arts in Anthropology and Archaeology from Hunter College, a City University of New York. Marina's graduate studies focused on social system dynamics in prehistoric Japan. She spent time doing research in Iceland and Japan before transitioning to Memorial Sloan Kettering Cancer Center in New York, where she worked as a Programmer/Analyst in Surgery Informatics. Most recently, Marina worked as a Data Architect at Stanford Health Care in California where she developed integrated data solutions for cancer quality initiatives. Marina is currently studying Natural Language Technology, an online certificate program with the University of Washington.

**GLENN ABASTILLAS**  
MA



**Glenn Abastillas** has joined the Surveillance Informatics Branch of SRP as an IMS Contractor doing Query Development. He works with the Natural Language Processing (NLP) team and clinical researchers on unstructured clinical data from state registries to creatively apply the latest advances in computational linguistic and NLP techniques.

Glenn is a computational linguist with research interests including investigating code-switching in digital contexts between two or more language systems. He speaks more than eight languages, including English, Cebuano, Spanish, German, French, Arabic, Portuguese, Norwegian, and Sign Language. He also 'speaks' several programming languages, including Python, Java, SQL and R. Glenn received his Bachelor of Science in Nursing from Rutgers, the State University of New Jersey and worked as a pediatric RN for children with epilepsy. He graduated with a Master of Arts in Linguistics from Georgetown University in Washington, D.C.

**MARIANNA PHILIP**  
MSc



**Marianna Philip** has joined the Administrative Support Team of SRP. She works with the administrative and communications teams.

Marianna received her Master's in Communication Sciences from Fatima Jinnah Women University in Pakistan. For the past 6 years, Marianna has provided administrative contract support at the National Center of Biotechnology Information in the National Library of Medicine. She has also worked as a contractor for NCI's Translational Research Program in the Division of Cancer Treatment and Diagnosis. Marianna is originally from Pakistan, where she attended a missionary school run by British nuns. She has been in the US for almost 9 years.

# UPDATES: NEW STAFF

**DONNA RIVERA**  
PharmD, MSc



**Donna Rivera** has joined the Surveillance Informatics Branch of SRP as a Scientific Project Officer. She works on projects that enhance the availability and quality of treatment data within SEER through various external collaborations, including evaluation of novel methodologies for treatment capture. She earned her Doctor of Pharmacy degree from the University of Florida College of Pharmacy in Gainesville, Florida and her Master of Science in Pharmaceutical Sciences from the University of Florida. She received her Bachelor of Science in Public Health from the University of South Florida.

She comes to SRP from the Clinical and Translational Epidemiology Branch (CTEB) of the Epidemiology and Genomics Research Program (EGRP) where she was a Program Director. Prior to NCI, Donna worked on Phase I clinical trials at Stiefel, a GlaxoSmithKline company.

**SRINI KRISHNAMOORTHY**  
PhD, MA



**Srini Krishnamoorthy** has joined SRP as a Cancer Research Training Award (CRTA) Fellow for communications. He works on projects and initiatives that help to improve the visibility of SEER programs and projects to public health workers, clinicians, researchers, policy makers, and the public.

Srini previously worked in academic and hospital-based cancer research settings. In his research career, Srini served in various roles such as Research Fellow, Post-Doctoral Researcher (Massachusetts General Hospital, Harvard Medical School), Principal Scientist at Physical Sciences Inc., and as a Research Professor at the University of Houston. He recently graduated from Emerson College with a Master of Arts in Health Communication.

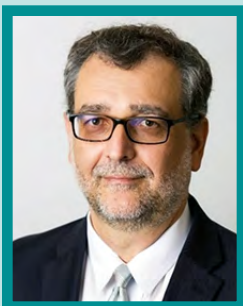
**GONÇALO FORJAZ DE LACERDA**  
DVM, MS



**Gonçalo Forjaz** has joined SRP as a Guest Researcher with research awards from Fulbright and Luso-American Development Foundation. Gonçalo has nine years of experience in the areas of cancer registration and descriptive epidemiology. Goncalo has worked on predictive analytics of the future cancer burden and the attributable causes of cancer in the Region of the Azores, Portugal. He has also provided guidance and training for NCI's work with the IARC's regional hub for the Caribbean countries. Gonçalo received his Doctorate of Veterinary Medicine from the University of Évora, Portugal and his Master of Science degree in Human Oncology from the University of Porto, Portugal.

As a consultant of IARC, Gonçalo participated in two impACT missions, one with the Republic of Angola and the other with the Republic of Mozambique, both of which were sponsored by the Programme of Action for Cancer Therapy of the International Atomic Energy Agency.

**SERBAN NEGOITA**  
MD, DrPH



**Serban Negoita** has joined SRP as Chief of the Data Quality, Analysis, and Interpretation Branch (DQAIB). He works with the SEER registries and SRP's experts to enhance the quality of cancer surveillance data. As Branch Chief, he facilitates DQAIB's mission to provide leadership in cancer data collection, quality improvement, education, analysis, interpretation, and reporting of the cancer burden.

Serban earned his MD from the University of Medicine and Pharmacy in Bucharest, Romania and his DrPH from the State University of New York. Serban's long-standing career includes many experiences worthy of note. While in Romania, he conducted surveillance of leukemia and lymphoma incidence for the International Agency for Research on Cancer (IARC) following the Chernobyl disaster. He also served at Maryland Cancer Registry where he and his colleagues achieved NAACCR gold certification four times. As an expert in cancer registry operations and data quality assessments, Serban anticipates that novel cancer technologies will require new quality assurance procedures and data quality measures.

# UPDATES: NEW STAFF

**SPENCER MORRIS**  
MS



**Spencer Morris** has joined SRP as a CRTA Fellow in the Surveillance Informatics Branch. He works with the Natural Language Processing (NLP) team on extracting clinical data from various forms of medical documents, enabling data integration to improve the utility of the SEER Program.

Spencer is originally from Columbus, Ohio. He earned a bachelor's degree in computer science and engineering from The Ohio State University in 2015. Spencer recently graduated from the University of Washington with a master's degree in computational linguistics. As an undergraduate intern at Philips Healthcare, he worked on calibration software for CT scanners. As a graduate intern at the Fred Hutchinson Cancer Research Center, he utilized NLP to create a system that automatically finds and extracts information about patients' tobacco and alcohol usage from free-text clinical notes, thereby reducing the time and cost of human abstractors to extract this data.

**ANA MEBANE**  
MA



**Ana Mebane** has joined the Communications Team of SRP as a CRTA Fellow. She helps advance the health communication efforts of SRP, such as projects related to social media and website content. Ana is originally from Chicago, IL. She earned her Master of Arts degree in Health Communication from DePaul University in Chicago and her Bachelor of Arts degree in Journalism from Howard University in Washington, D.C.

Prior to joining NCI, Ana worked in patient registration for Northwestern Medicine. She was also an Assessment Support Assistant for DePaul University's Office of Teaching and Learning Assessment, and wrote health-related blogs for several hospitals.

**ALYSSA WANG**  
MPH



**Alyssa Wang** has joined the Communications Team of SRP as a CRTA Fellow. She works on various projects related to generating communication materials such as infographics and newsletters. She also helps coordinate activities for the SEER Data Management System (SEER\*DMS) and the SEER-linked Virtual Tissue Repository Pilot Study.

Alyssa is originally from Ponte Vedra Beach, Florida. She earned dual bachelor's degrees in Microbiology and Public Relations and a Master of Public Health degree from the University of Florida. Most recently, Alyssa served as a legislative correspondent in the U.S. Senate on Capitol Hill, where she worked on federal health policy issues and communicated with constituents on various health topics. Alyssa has also interned with a research team at the Department of Veterans Affairs in Gainesville, Florida, to evaluate a program for caregivers.

**MICHAEL APATA**  
MPH



**Michael Apata** is an At-Large Presidential Management Fellow, enabling him to rotate at various institutes within NIH. He has joined SRP to provide project management support to multiple programs. He is involved in the NCI's collaboration with the Department of Energy (DOE), Data Acquisition/Linkage effort, and other projects as needed. He also provides contract management support for the SEER renewal.

Prior to joining SRP, Michael completed a rotation at the NIH Office of Research Services where he worked on its Workforce Enhancement Initiative and also completed rotations at the Fogarty International Center and CDC Washington Office where he worked on communication and legislative issues. Michael earned a bachelor's degree in accounting and finance from Nigeria and a master's degree in public health from Texas A&M University, concentrating in health policy and management.

# UPDATES: NEW STAFF

**MITA MYNENI**  
MBA, MS



**Mita Myneni** has joined SRP as a Center for Biomedical Informatics and Information Technology (CBIIT) contractor through NetImpact Strategies, Inc. She provides project management and content management support on Pilot 3 (Population Level Pilot for Population Information Integration, Analysis, and Modeling) of the Joint Design of Advanced Computing Solutions for Cancer (JDACS4C) program, NCI's collaboration with the DOE.

Mita is originally from India and moved to the U.S. to pursue her education. She earned her undergraduate degree in Computer Science and an MBA focusing on International Business & Marketing. Most recently, she earned an MS in Marketing and Research from the Zicklin School of Business at Baruch College in New York City. Before joining NCI's SRP team, Mita worked in the media, retail, and oil industries in the areas of business development, marketing, and project management.

**USMAN KHALID**



**Usman Khalid** has joined SRP as a CBIIT contractor in the Surveillance Informatics Branch (SIB). He works with the SIB team on Pilot 3 of the NCI-DOE Collaboration, which will deliver an infrastructure that will support the development of algorithms and informatics tools to enable a comprehensive, scalable, and cost-effective national cancer surveillance program.

Usman was born in Pakistan and immigrated to Northern Virginia as a toddler. He earned a Bachelor of Science in biology from Virginia Commonwealth University. After graduating, Usman worked for pharmaceutical company GlaxoSmithKline for about a year. He then made a career transition into healthcare IT, where he worked as an Application Analyst at Evolent Health.

**ANDREW GROTHEN**  
PharmD



**Andrew Grothen** has joined the Surveillance Informatics Branch of SRP as a CRTA Fellow for Cancer Treatment Data Linkage. Andrew is part of the Data Acquisition and Linkages Team to identify methods and data that more efficiently capture detailed information on treatment for SEER.

Andrew is originally from the state of New Hampshire. He earned his Doctor of Pharmacy degree from the University of Wyoming School of Pharmacy in Laramie, Wyoming. He is a licensed pharmacist and was working at Safeway Pharmacy in Nebraska prior to joining SRP. At Safeway, he performed various duties including dispensing prescription medication, providing patient counseling, performing medication therapy management, and training technicians. He has also completed internships and rotations in pharmacy in public and private settings, including the U.S. Department of Veterans Affairs, the Indian Health Service, and nonprofit hospitals. Andrew has conducted various research projects, including examining the use of a specific diabetes medication in a community setting, antibiotic use for bronchitis, and patient education for asthma and chronic obstructive pulmonary disease.

# INITIATIVES

## • NCI-DOE Collaboration

The SEER Program is collaborating with four national laboratories of the Department of Energy (DOE) on a pilot project that will deliver an infrastructure that will support the development of algorithms and informatics tools to enable a comprehensive, scalable, and cost-effective national cancer surveillance program that enhances the existing system while also expanding the breadth of data captured to integrate biological, social, psychological and ecological variables to model for cancer outcomes.



### The main aims of the project are:

**Aim 1:** Develop scalable machine-learned natural language processing (NLP) tools for deep comprehension of unstructured clinical text to enable automated and accurate capture of reportable cancer surveillance data elements with case-specific uncertainties quantified to guide manual abstraction.

**Aim 2:** Build a working infrastructure that integrates large scale linkage of heterogeneous data sources for research and discovery with a focus on methodology and bias assessment. Scalable graph and visual analytics tools guide exploration of disease specific cancer surveillance algorithms and methods to accelerate understanding of cancer outcomes including progression, recurrence, metastases, and survival.

**Aim 3:** Develop a cancer surveillance data-driven modeling and simulation environment for predicting health trajectories of cancer patients.

SRP, in collaboration with IMS, LabKey, and Linguamatics, has developed an infrastructure for document selection, annotation, algorithm development and data hosting which will be utilized for human annotation of data elements in e-path reports to facilitate algorithm development. Using deep learning methods, DOE researchers have been developing robust NLP tools that will identify, with known uncertainty, site, histology, laterality and behavior in pathology reports that will be utilized by cancer registries. External heterogeneous data sources have been identified for linkage with SEER data to enable development of longitudinal patient trajectories that will support modeling efforts. Extensive work in variable selection, clinical treatment pathways, and algorithm development is ongoing. Furthermore, clinical experts have joined the collaboration to identify and develop modeling use cases for specific cancer sites.

Read more about it in our [blog post](#).

# NEW ON THE 'NET

## • SRP Blog

**SRP has launched a new blog series. In this blog, we share updates with our many stakeholders about initiatives that SRP is spearheading to bring the nation closer to a complete and detailed understanding of the cancer burden.**

The [blog's](#) commentaries report on key collaborations and other efforts that aim to build an increasingly powerful and useful cancer surveillance infrastructure. We also highlight methods and technologies that are and will be essential for rapidly collecting, analyzing, interpreting, sharing, and applying cancer surveillance data. This blog reflects SRP's research priorities, and is intended to further SRP's mission. In view of the national Precision Medicine Initiative and Cancer Moonshot, this is a critical time for the cancer community to focus on precision cancer surveillance.



## • Cancer Trends Progress Report

*Online Summary of US Cancer Control Measures | [Progressreport.cancer.gov](http://Progressreport.cancer.gov)*

**As part of a collaboration with the Implementation Science team, the Surveillance Research Program released the latest Cancer Trends Progress Report in January this year.**

Since 2001, this report has been a key source of national trend data across a spectrum of cancer control measures. From prevention to end of life, the site offers trends by sex, age, race, ethnicity, and income and education level, and displays them in relation to Healthy People 2020 targets when available. The Cancer Trends Progress Report can be used by researchers and cancer control professionals to elicit research ideas and set priorities for cancer control program planning to advance cancer control progress.

 **NATIONAL CANCER INSTITUTE**  
Cancer Trends Progress Report



The Cancer Trends Progress Report gathers data from several federal departments and agencies, such as the Environmental Protection Agency, the Federal Trade Commission, the Department of Agriculture, and many offices and agencies within the Department of Health and Human Services (including the Centers for Disease Control and Prevention, the Office of Disease Prevention and Health Promotion, the Substance Abuse and Mental Health Administration, and the National Institute on Alcohol Abuse and Alcoholism). This report is developed with input from federal partners, consumer advocates, non-profits, and others.

“We at NCI, along with our Cancer Trends Progress Report partners, hope that you will find this report to be a valuable reference tool and a stimulus for action. We must not forget that the numbers in this report reflect the lives and struggles of millions of our fellow citizens. NCI remains committed to advancing scientific progress and facilitating its application on behalf of each of them.” –*Bob Croyle, Director, Division of Cancer Control and Population Sciences*

Visit the [site](#) to view measures and generate custom reports.



# NEW ON THE 'NET continued

## • Annual Report to the Nation (ARN) *Special Section: Survival*

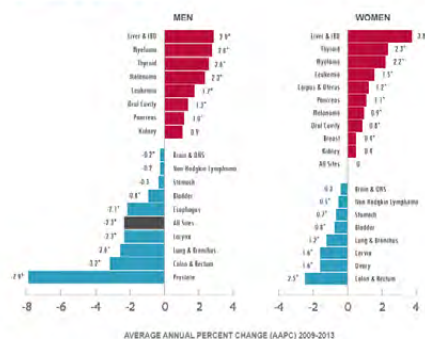
**The Annual Report to the Nation on the Status of Cancer, 1975-2014, is an update of rates for new cases and deaths as well as trends for the most common cancers in the United States. The special section of this report focuses on survival.**

Since our last newsletter, the Surveillance Research Program has collaborated with the North American Association of Central Cancer Registries (NAACCR), the American Cancer Society (ACS), and the Centers for Disease Control and Prevention (CDC) on the Annual Report to the Nation on the Status of Cancer, 1975-2014 (ARN). The report provided an overview of cancer trends as well as a special section on survival by stage, race/ethnicity, and state for common cancers. This year's ARN highlighted continuing declines in cancer death rates for men, women, and children, and for nearly all major cancer sites. From 2000 to 2014, death rates declined by 1.8 percent per year among men and 1.4 percent per year among women, and 1.7 percent per year for people ages 0 to 19. The incidence report highlighted that new cases of lung cancer have continued to decrease along with tobacco use. It is noted that overall cancer incidence rates continued to decrease among men, remained stable among women, and increased among children.

This year's special section emphasized that overall survival rates from 1975 to 2012 improved for the most common cancers, even when patients received diagnoses of distant-stage disease. However, cancer survival rates can vary based on race, ethnicity, and state. Differences in survival by race that existed in 1975 are still evident in the data from 2012. This is largely due to biological differences, socioeconomic status, and access to health care. The report also highlighted survival of childhood cancers, which improved from 58% to 85%. Researchers attribute many improvements in childhood cancer survival to better supportive care and clinical trials of combination chemotherapy, radiotherapy, and/or surgery.

To read the full report or see previous Annual Reports to the Nation, visit [https://seer.cancer.gov/report\\_to\\_nation](https://seer.cancer.gov/report_to_nation)

NATIONAL INCIDENCE TRENDS



\*AAPC is significantly different from zero ( $p < .05$ ). Rates were adjusted for reporting delay in the registry.

seer.cancer.gov  
Source: Annual Report to the Nation on the Status of Cancer 1975-2014



# NEW ON THE 'NET continued

## • SEER\*Stat

SRP has recently released a new version of SEER\*Stat (Version 8.3.4).

The download and installation information for the new version can be found here: <https://seer.cancer.gov/seerstat>.



## • NCI Map Stories

NCI Map Stories are GIS Portal-based interactive maps that help in visualization of cancer related geo-spatial data.

These maps are updated periodically to reflect new data and educate the user about a specific cancer topic. See our new updated Map Stories for lung, breast, and prostate cancer here: <https://gis.cancer.gov/mapstory>.



## • SEER\*Explorer

SEER\*Explorer is a new interactive website to provide easy access to a wide range of SEER cancer statistics. It provides detailed statistics for a cancer site by gender, race, calendar year, age, and for a selected number of cancer sites, by stage and histology.



With SEER\*Explorer, you can:

- Create custom graphs and tables
- Download data and images
- Download SEER\*Stat sessions
- Share links to results

The initial release of SEER\*Explorer is a beta version that provides access to SEER incidence and U.S. mortality statistics, including:

- Recent Trends, 2000-2013
- Long Term Trends, 1975-2013
- Rates by Age
- Stage Distribution (for incidence data only)

You can test out the beta version here: <http://seer.cancer.gov/explorer/>

# NEW ON THE 'NET continued

## • Update on the Did You Know? Video Series

The Did You Know? Video Series provides 3- to 4- minute informational videos on various cancer topics. The videos communicate key statistical data on different types of cancer to a lay audience in understandable, clear, and concise language. The Did You Know? Team has recently released videos on [Oral Cavity & Pharynx Cancer](#), [Hepatitis & Liver Cancer](#), [Kidney & Renal Pelvis Cancer](#), [Endometrial Cancer](#), [Brain & Other Nervous System Cancers](#), and [Ovarian Cancer](#).

Did You Know? Video Series Triumphs: Nine Did You Know? Videos are in NCI's top 20. Our video on [Human Papillomavirus \(HPV\)](#) has more than 25,000 views, the [Cancer Statistics Video](#) has over 17,000 views, and the [Leukemia Statistics Video](#) has more than 15,000 views! Additionally, the videos on [Breast Cancer Statistics](#) and [Lung Cancer](#) have more than 10,000 views each.

This year, the team looks forward to adding more cancer-related topics and is currently developing videos in Spanish to be featured on [cancer.gov/espanol](http://cancer.gov/espanol). These videos are in the public domain, free and available for your use. Get the conversation started by sharing on social media. Embed them on any website or presentation or email them to family and friends. You can find them on our SEER website at <https://seer.cancer.gov/statistics/videos/>.



## • @NCICancerStats reaches 5,000 Twitter followers! Help us continue to grow!

In June 2017, the @NCICancerStats Twitter handle reached 5,000 followers.

[@NCICancerStats](#) provides the latest cancer statistics, information on new online tools, and resources for researchers. Please follow us, retweet us, or like our tweets if you haven't already!



# DATA

## • SEER Releases New Data

**SEER released new data on April 14, 2017, which also includes the new Cancer Statistics Review (CSR).**

This dataset includes SEER incidence and population data associated by age, sex, race, year of diagnosis, and geographic area. In 2014, there were 416,914 malignant cases reported, with a total of 8,662,369 malignant cases recorded from 1973–2014. For more information, visit <http://seer.cancer.gov/data/>.

Since the early 1970s, the Surveillance, Epidemiology, and End Results (SEER) Program has been an invaluable resource for statistics on cancer in the United States, tracking and reporting trends in incidence, mortality, survival, and prevalence. Researchers at NCI and around the country continue to rely on SEER for the most accurate cancer statistics.

	All Cases*	Malignant Cases	Malignant + In Situ Cases
1973-2014 Data (Nov 2016 Submission)	9,675,661	8,562,369	9,429,370
1973-2013 Data (Nov 2015 Submission)	9,176,963	8,234,845	8,958,804
Increase from 2015 to 2016 Submission	498,698	442,724	470,575
Number of 2014 cases	482,000	416,914	459,473

# EVENTS

## • Summary of NLP Workshop

On December 8-9, 2016, the National Cancer Institute (NCI) hosted a joint workshop with the Centers for Disease Control and Prevention (CDC) and Food and Drug Administration (FDA) entitled Natural Language Processing (NLP) & Machine Learning for Cancer Surveillance. The objective of this meeting was to identify best approaches and methods to extract, classify, and consolidate structured data from heterogeneous sources of free-text documents. The workshop featured presentations that covered a variety of topics including summarizing experiences with NLP tools, NLP research design, tool evaluation methods, and approaches for de-identification. Over 85 people from government, academia, industry, and cancer registries participated in-person or via WebEx to discuss successes and challenges with natural language processing.

## • SEER\*DMS Face-to-Face Meeting

*July 12-14, 2017*



The Surveillance, Epidemiology, and End Results (SEER) Program of NCI is actively working to enhance the data collected for cancer surveillance with the goal of providing richer datasets for the cancer research community. SEER cancer registries are the bedrock of these initiatives, and one of the main goals of SEER is to ensure that we can support the development and enhancement of the right tools and processes for the registry community.

The SEER Program is evolving and will be experiencing changes in automation, natural language processing, linkages, quality initiatives, and more. SEER is working with Information Management Systems, Inc. (IMS) to evaluate and enhance the SEER Data Management System (SEER\*DMS), actively engage the registry community to better understand their fundamental and evolving needs, and develop new processes for future developments within the SEER\*DMS infrastructure.

The SEER\*DMS Face-to-Face Meeting will be held July 12-14, 2017, at the NCI Shady Grove Campus in Rockville, Maryland. The meeting will allow the SEER Program, registries, and IMS to work together to discuss common issues to progress forward as a SEER community. The meeting will also bring awareness to the current changes happening in SEER\*DMS, highlight upcoming initiatives for the platform, and gather functional requirements as a SEER\*DMS community to enhance the system.

# PUBLICATIONS

## • SRP Staff Publications

**BMJ  
Open**

**Cancer  
Epidemiology,  
Biomarkers  
& Prevention**

**Cancer Causes  
& Control**

**AJG**

**International  
Journal of Health  
Geographics**

**Cancer**

**Adamo MP, Boten JA, Coyle LM, Cronin KA, Lam CJ, Negoita S, Penberthy L, Stevens JL, Ward KC.** Validation of prostate-specific antigen laboratory values recorded in Surveillance, Epidemiology, and End Results registries. *Cancer*. 2017 Feb 15;123(4):697-703. doi: 10.1002/cncr.30401. [[PubMed Abstract](#)]

**Altekruse S, Das A, Cho H, Petkov VI, Yu M.** Do US thyroid cancer incidence rates increase with socioeconomic status among people with health insurance? An observational study using SEER population-based data. *BMJ Open* 2015;5(12). doi: 10.1136/bmjopen-2015-009843. [[PubMed Abstract](#)]

**Andridge R, Noone AM, Howlader N.** Imputing estrogen receptor (ER) status in a population-based cancer registry: a sensitivity analysis. *Stat Med*. 2017 Mar 15;36(6):1014-1028. doi: 10.1002/sim.7193. [[PubMed Abstract](#)]

**Barr RD, Ferrari A, Ries L, Whelan J, Bleyer WA.** Cancer in adolescents and young adults: a narrative review of the current status and a view of the future. *JAMA Pediatr*. 2016 May 1;170(5):495-501. doi: 10.1001/jamapediatrics.2015.4689. [[PubMed Abstract](#)]

**Barr RD, Ries LA, Lewis DR, Harlan LC, Keegan TH, Pollock BH, Bleyer WA.** Incidence and incidence trends of the most frequent cancers in adolescent and young adult Americans, including “nonmalignant/noninvasive” tumors. *Cancer*. 2016 Apr 1;122(7):1000-8. doi: 10.1002/cncr.29867. [[PubMed Abstract](#)]

**Blecker S, Johnson NJ, Altekruse S, Horwitz LI.** Association of occupation as a physician with likelihood of dying in a hospital. *JAMA*. 2016 Jan 19;315(3):301-3. doi: 10.1001/jama.2015.16976. [[Full Text Link](#)]

**Bluethmann SM, Mariotto AB, Rowland JH.** Anticipating the “silver tsunami”: prevalence trajectories and comorbidity burden among older cancer survivors in the United States. *Cancer Epidemiol Biomarkers Prev* 2016;25(7):1029-1036. doi: 10.1158/1055-9965.EPI-16-0133. [[PubMed Abstract](#)]

# PUBLICATIONS

## Publications continued...

Breen N, **Lewis DR**, Gibson JT, Yu M, Harper S. Assessing disparities in colorectal cancer mortality by socioeconomic status using new tools: health disparities calculator and socioeconomic quintiles. *Cancer Causes Control*. 2017 Feb;28(2):117-125. doi: 10.1007/s10552-016-0842-2. [\[PubMed Abstract\]](#)

Duggan MA, Anderson WF, **Altekruse S**, **Penberthy L**, Sherman ME. The Surveillance, Epidemiology, and End Results (SEER) Program and pathology: toward strengthening the critical relationship. *Am J Surg Pathol* 2016;40(12). doi: 10.1097/PAS.0000000000000749. [\[PubMed Abstract\]](#)

**Gonzales FA**, **Taplin SH**, **Yu M**, Breen N, **Cronin KA**. Receipt of mammography recommendations among White and non-White women before and after the 2009 United States Preventive Services Task Force recommendation change. *Cancer Causes Control* 2016;27(8):977-987. doi: 10.1007/s10552-016-0775-9. [\[PubMed Abstract\]](#)

Gulati R, **Feuer EJ**, Etzioni R. Conditions for valid empirical estimates of cancer overdiagnosis in randomized trials and population studies. *Am J Epidemiol* 2016;184(2):140-147. doi: 10.1093/aje/kwv342. [\[PubMed Abstract\]](#)

Gupta S, Aitken JF, Bartels U, Brierley J, Dolendo M, Friedrich P, Fuentes-Alabi S, Garrido CP, Gatta G, Gospodarowicz M, Gross T, Howard SC, Molyneux E, Moreno F, Pole JD, Pritchard-Jones K, Ramirez O, **Ries LA**, Rodriguez-Galindo C, Shin HY, Steliarova-Foucher E, Sung L, Supriyadi E, Swaminathan R, Torode J, Vora T, Kutluk T, Frazier AL. Paediatric cancer stage in population-based cancer registries: the Toronto consensus principles and guidelines. *Lancet Oncol*. 2016 Apr;17(4):e163-72. doi: 10.1016/S1470-2045(15)00539-2. pii: S1470-2045(15)00539-2. [\[PubMed Abstract\]](#)

Han J, **Zhu L**, Kulldorff M, Hostovich S, Stinchcomb DG, **Tatalovich Z**, **Lewis DR**, **Feuer EJ**. Using Gini coefficient to determining optimal cluster reporting sizes for spatial scan statistics. *Int J Health Geogr* 2016;15(1):27. doi: 10.1186/s12942-016-0056-6. [\[PubMed Abstract\]](#)

Han SS, Ten Haaf K, Hazelton WD, Munshi VN, Jeon J, Erdogan SA, Johanson C, McMahon PM, Meza R, Kong CY, **Feuer EJ**, de Koning HJ, Plevritis SK. The impact of overdiagnosis on the selection of efficient lung cancer screening strategies. *Int J Cancer*. 2017 Jun 01;140(11):2436-2443. doi: 10.1002/ijc.30602. [\[PubMed Abstract\]](#)

**Howlader N**, Morton LM, **Feuer EJ**, Besson C, Engels EA. Contributions of subtypes of non-Hodgkin lymphoma to mortality trends. *Cancer Epidemiol Biomarkers Prev*. 2016 Jan;25(1):174-9. doi: 10.1158/1055-9965.EPI-15-0921. [\[PubMed Abstract\]](#)

# PUBLICATIONS



## Publications continued...

**Howlader N**, Shiels MS, **Mariotto AB**, Engels EA. Contributions of HIV to non-Hodgkin lymphoma mortality trends in the United States. *Cancer Epidemiol Biomarkers Prev* 2016;25(9):1289-1296. doi: 10.1158/1055-9965.EPI-16-0273. [\[PubMed Abstract\]](#)

Jemal A, Ward EM, Johnson CJ, **Cronin KA**, Ma J, Ryerson B, **Mariotto A**, Lake AJ, Wilson R, Sherman RL, Anderson RN, Henley SJ, Kohler BA, **Penberthy L**, **Feuer EJ**, Weir HK. Annual Report to the Nation on the Status of Cancer, 1975-2014, featuring survival. *J Natl Cancer Inst.* 2017 Sep 01;109(9)doi: 10.1093/jnci/djx030. [\[PubMed Abstract\]](#)

Keegan TH, **Ries LA**, Barr RD, Geiger AM, Dahlke DV, Pollock BH, Bleyer WA, National Cancer Institute Next Steps for Adolescent and Young Adult Oncology Epidemiology Working Group. Comparison of cancer survival trends in the United States of adolescents and young adults with those in children and older adults. *Cancer* 2016;122(7):1009-1016. doi: 10.1002/cncr.29869. [\[PubMed Abstract\]](#)

Kish JE, Rolin AI, Zou Z, Cucinelli JE, **Tatalovich Z**, Saraiya M, **Altekruse SF**. Prioritizing US cervical cancer prevention with results from a geospatial model| *Global Oncol* 2016;2(5):275-283. doi:10.1200/JGO.2015.001677. [\[Link to Article\]](#)

Lee M, Gouskova NA, **Feuer EJ**, Fine JP. On the choice of time scales in competing risks predictions. *Biostatistics.* 2017 Jan;18(1):15-31. doi: 10.1093/biostatistics/kxw024. [\[PubMed Abstract\]](#)

**Lewis DR**, **Chen HS**, Cockburn M, Wu XC, Stroup AM, Midthune DN, Krapcho MF, Miller DG, et al. Preliminary estimates of SEER cancer incidence for 2013. *Cancer* 2016;122(10):1579-1587. doi: 10.1002/cncr.29953. [\[PubMed Abstract\]](#)

**Liu B**, **Yu M**, Graubard BI, Troiano RP, Schenker N. Multiple imputation of completely missing repeated measures data within person from a complex sample: application to accelerometer data in the National Health and Nutrition Examination Survey. *Stat Med* 2016;35(28):5170-5188. doi: 10.1002/sim.7049. [\[PubMed Abstract\]](#)

Major JM, **Penberthy L**, McGlynn KA. Data systems and record linkage: considerations for pharmacoepidemiologic studies examining cancer risk. *Ann Epidemiol.* 2016 Nov;26(11):746-748. pii: S1047-2797(16)30279-4. doi: 10.1016/j.annepidem.2016.08.015. [\[Full Text Link\]](#)





# PUBLICATIONS

## Publications continued...

Mandelblatt JS, Stout NK, Schechter CB, van den Broek JJ, Miglioretti DL, Krapcho M, Trentham-Dietz A, Munoz D, Lee SJ, Berry DA, van Ravesteyn NT, Alagoz O, Kerlikowske K, Tosteson AN, Near AM, Hoeffken A, Chang Y, Heijnsdijk EA, Chisholm G, Huang X, Huang H, Ergun MA, Gangnon R, Sprague BL, Plevritis S, **Feuer EJ**, de Koning HJ, **Cronin KA**. Collaborative modeling of the benefits and harms associated with different U.S. breast cancer screening strategies. *Ann Intern Med*. 2016 Feb 16;164(4):215-25. doi: 10.7326/M15-1536.

[\[PubMed Abstract\]](#)

**Mariotto A**, Etzioni R, Hurlbert M, **Penberthy L**, and Mayer M. Estimation of the number of women living with metastatic breast cancer in the United States. *Cancer Epidemiology Biomarkers Prev*. Online May 18, 2017. In print June 1, 2017. DOI: 10.1158/1055-9965.EPI-16-0889. [\[Link to Article\]](#)

Miller KD, Siegel RL, Lin CC, **Mariotto AB**, Kramer JL, Rowland JH, Stein KD, Alteri R and Jemal A. Cancer treatment and survivorship statistics, 2016. *CA Cancer J Clin* 2016;66(4):271-289. doi: 10.3322/caac.21349.

[\[PubMed Abstract\]](#)

**Noone AM**, **Cronin KA**, **Altekruse SF**, **Howlander N**, **Lewis DR**, **Petkov VI**, **Penberthy L**. Cancer incidence and survival trends by subtype using data from the Surveillance Epidemiology and End Results Program, 1992-2013. *Cancer Epidemiol Biomarkers Prev*. 2017 Apr;26(4):632-641. doi: 10.1158/1055-9965.EPI-16-0520.

[\[PubMed Abstract\]](#)

Ostrom QT, Gittleman H, Kruchko C, Louis DN, Brat DJ, Gilbert MR, **Petkov VI**, Barnholtz-Sloan JS. Completeness of required site-specific factors for brain and CNS tumors in the Surveillance, Epidemiology and End Results (SEER) 18 database (2004-2012, varying). *J Neurooncol* 2016;130(1):31-42. doi: 10.1007/s11060-016-2217-7. [\[PubMed Abstract\]](#)

**Petkov VI**, Miller DP, **Howlander N**, Gliner N, Howe W, Schussler N, **Cronin KA**, Baehner FL, Cress R, Deapen D, Glaser SL, Hernandez BY, Lynch CF, Mueller L, Schwartz AG, Schwartz SM, Stroup A, Sweeney C, Tucker TC, Ward KC, Wiggins C, Wu X, **Penberthy L** & Shak S. Breast-cancer-specific mortality in patients treated based on the 21-gene assay: a SEER population-based study. *NPJ Breast Cancer* 2016;2():16017. doi: 10.1038/nnpjbcancer.2016.17. [\[PubMed Abstract\]](#)

**Ries L**, Trama A, Nakata K, Gatta G, Botta L, Bleyer A. Cancer incidence, survival, and mortality among adolescents and young adults. In: Bleyer A, Barr R, Ries L, Whelan J, Ferrari A (eds.). *Cancer in Adolescents and Young Adults*, 2nd Edition. Springer Publishing; 2017; p.7-42. [\[Abstract\]](#)

# PUBLICATIONS

## Publications continued...



Roberts MC, Miller DP, Shak S, **Petkov VI**. Breast cancer-specific survival in patients with lymph node-positive hormone receptor-positive invasive breast cancer and Oncotype DX Recurrence Score results in the SEER database. *Breast Cancer Res Treat*. 2017 Jun;163(2):303-310. doi: 10.1007/s10549-017-4162-3. [[PubMed Abstract](#)]

Ryerson AB, Ehemann CR, **Altekruse SF**, Ward JW, Jemal A, Sherman RL, Henley SJ, Holtzman D, Lake A, **Noone AM**, Anderson RN, Ma J4, Ly KN, **Cronin KA**, **Penberthy L**, Kohler BA. Annual Report to the Nation on the Status of Cancer, 1975-2012, featuring the increasing incidence of liver cancer. *Cancer* 2016;122(9):1312-1337. doi: 10.1002/cncr.29936. [[PubMed Abstract](#)]

**Serrano KJ**, **Yu M**, Coa KI, Collins LM, Atienza AA. Mining health app data to find more and less successful weight loss subgroups. *J Med Internet Res* 2016;18(6). doi: 10.2196/jmir.5473. [[PubMed Abstract](#)]

Shen X, Wang L, **Zhu L**. Spatial analysis of regional factors and lung cancer mortality in China, 1973-2013. *Cancer Epidemiol Biomarkers Prev*. 2017 Apr;26(4):569-577. doi: 10.1158/1055-9965.EPI-16-0922. [[PubMed Abstract](#)]

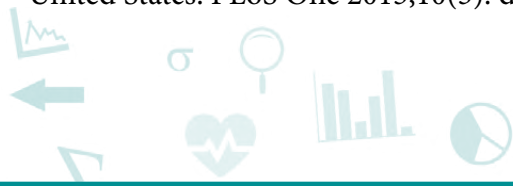
**Steeves JA**, **Liu B**, Willis G, Lee R, Smith AW. Physicians' personal beliefs about weight-related care and their associations with care delivery: the U.S. National Survey of Energy Balance Related Care among Primary Care Physicians. *Obes Res Clin Pract* 2015;9(3):243-255. doi: 10.1016/j.orcp.2014.08.002. [[PubMed Abstract](#)]

Temkin SM, Minasian L, **Noone AM**. The end of the hysterectomy epidemic and endometrial cancer incidence: what are the unintended consequences of declining hysterectomy rates?. *Front Oncol* 2016; 6:89. doi: 10.3389/fonc.2016.00089. [[PubMed Abstract](#)]

Ten Haaf K, Jeon J, Tammemägi MC, Han SS, Kong CY, Plevritis SK, **Feuer EJ**, de Koning HJ, Steyerberg EW, Meza R. Risk prediction models for selection of lung cancer screening candidates: a retrospective validation study. *PLoS Med*. 2017 Apr;14(4):e1002277. doi: 10.1371/journal.pmed.1002277. [[PubMed Abstract](#)]

Thomas A, Link BK, **Altekruse S**, Romitti PA, Schroeder MC. Primary breast lymphoma in the United States: 1975-2013. *J Natl Cancer Inst*. 2017 Jun 01;109(6)doi: 10.1093/jnci/djw294. [[PubMed Abstract](#)]

Zou Z, McGlynn KA, **Tatalovich Z**, **Altekruse SF**, Petrick JL, **Rolin AI**, Cuccinelli JE. Geographic variation of intrahepatic cholangiocarcinoma, extrahepatic cholangiocarcinoma, and hepatocellular carcinoma in the United States. *PLoS One* 2015;10(3). doi: 10.1371/journal.pone.0120574. [[PubMed Abstract](#)]



# GRANTS

## • SRP Grants Awarded in Fiscal Year 2016

*Newly funded SRP competing grant awardees for Fiscal Year 2016 are listed below. In addition to these newly funded grants, SRP received and reviewed over 80 new grant applications and currently manages about 100 existing, non-competing grants which were processed for continued funding.*

**KEY**

**DAB** – Data Analytics Branch

**DQAIB** – Data Quality Analysis and Interpretation Branch

**OAD** – Office of the Associate Director

**SRAB** – Statistical Research and Applications Branch

SRP Branch	Program Director	Principal Investigator	Research Project Title	Institution
DQAIB	Denise Lewis	Sudha Xirasagar	Reducing Colorectal Cancer Disparities: Racial Differences in Colorectal Polyp Profile	University of South Carolina at Columbia
OAD	Rose Fredua	Betsy Kohler	2016 North American Association of Central Cancer Registries Annual Conference	North American Association of Central Cancer Registries
OAD	Steve Friedman	Hua Xu	Advancing Cancer Pharmacoepidemiology Research Through EHRs and Informatics	University of Texas Health Science Center Houston
DAB	Angela Mariotto	Keiko Akagi	Dr. Keiko Akagi, A Bioinformatics Specialist in Cancer Genomics Research at OSU Comprehensive Cancer Center	Ohio State University
DAB	Angela Mariotto	Brian L. Egleston	Deep Learning for Representation of Codes Used for SEER-Medicare Claims Research	Research Institute of Fox Chase Cancer Center
DAB	Angela Mariotto	Liang Zhu	New Methods to Address Dilemmas in Mixed Recurrent-Event and Panel-Count Data	University of Texas Health Science Center Houston
DAB	Angela Mariotto	Jing Ning	Comparative Effectiveness of Cancer Research: Use Data from Multiple Sources	University of Texas MD Anderson Cancer Center
DAB	Angela Mariotto	Carmen Lynn Lewis	Improving Targeted Colorectal Cancer Screening in the Elderly	University of Colorado Denver
SRAB	Huann-Sheng Chen	Jaya Satagopan	Statistical and Computational Methods for Pharmacogenetic Epidemiology of Cancer	Memorial Sloan-Kettering Institute of Cancer Research
SRAB	Huann-Sheng Chen	Jaya Satagopan	Study of Exposures and Biomarkers in Cancer Epidemiology	Memorial Sloan-Kettering Institute of Cancer Research
SRAB	Huann-Sheng Chen	Valen Earl Johnson	Consistent Variable Selection in $p \gg n$ Settings	Texas A&M University
SRAB	Benmei Liu	Stuart R. Lipsitz	Analyzing Complex Cancer Studies with Skewed Responses	Brigham and Women's Hospital



 Follow us @NCICancerStats