Enhancements to NCI's SEER Program Creating New Research Opportunities

Forty-five years since its launch, SEER can become an invaluable resource for performing and informing clinical and even basic research.

September 6, 2018 By National Cancer Institute

By Lynne Penberthy, MD

When NCI launched the Surveillance, Epidemiology, and End Results (SEER) program 45 years ago, few could have anticipated that one day it would rely on powerful supercomputers to scan electronic pathology reports to collect real-time information about patients' tumors.

And, yet, this is exactly what is happening in one of a series of innovative SEER pilot studies that represent a fundamental rethinking of what this long-running program can and should be. These pilot studies are setting the stage for SEER to routinely collect far more clinically relevant pieces of data at the population level than previously—data that the research community as well as patients want and need to better understand cancer and its impact.

Although many of these new data collection efforts are in their early stages, they represent a true sea change for SEER. We should no longer consider SEER as just a tool for conducting population-based observational studies. I believe it can also become an invaluable resource for performing and informing clinical and even basic research.

More Data on Treatments, Genomics, and Recurrences

When we set about to enhance SEER, we worked with and sought feedback from many groups and experts. This outreach allowed us to better understand the cancer research community's needs and the opportunities that could be created through strategic, targeted enhancements to SEER.

In those discussions and forums, the main data needs that were identified included those that cover:

- All treatments received by individual patients from diagnosis until their death
- Outcomes other than incidence and mortality—in particular, recurrences
- The genomic composition of patients' tumors

As might be expected, enabling SEER to collect these data is complicated. Doing so requires capturing information over an extended time period from a massive, fragmented US health care system that relies on a wide variety of records systems that are not designed for easy information sharing.

Take, for example, a woman with early-stage breast cancer diagnosed and successfully treated in 2010 but who then had a recurrence in 2015. Ideally, we want to capture data concerning her primary treatment (e.g., surgery, radiation); any prescribed adjuvant treatments (e.g., chemotherapy, hormonal therapy) and whether they were actually administered; important clinical factors (e.g., non-cancer health problems, age, race/ethnicity); where in the body her recurrence occurred and how it was diagnosed; the outcomes of germline and tumor genomic tests done at the time of recurrence; and what treatments she received for that recurrence.

Some of these data may be in free-form notes (general notes for which there are no specific data fields) in an electronic record. Some may be in pharmacy records, or a primary care physician's notes, or insurance records. And if the woman moved to a different state during this period, that presents a whole new layer of challenges.

Trying to collect such data, and ensure that they are accurate, is an extremely complex undertaking. It's because of that great complexity that we are moving forward with these enhancements in a thoughtful and systematic way—one that we believe will lay the foundation for successfully integrating them into the broader SEER program.

More Cases, More Diversity

Among the most important enhancements to the SEER program is the expansion of the size and diversity of the population it covers.

SEER now covers 34% of the US population, covering 19 different geographic areas. With this expansion, SEER is collecting data on approximately 550,000 new individual cancer cases (without any personally identifying information) each year.

This expansion was implemented with a strategic focus: adding US cancer registries that include more underserved and ethnic/racial minorities. That means SEER is capturing data on cancer that more fully represents the US population.

Of course, it's always good to have more data to work from. In practice, the expanded population covered by SEER means that researchers will be able to perform better studies—those that allow us to more completely understand how cancer affects different patient subgroups and inform the development of interventions intended to address the shortcomings and disparities that we know exist in the current system of cancer care.

Pilot Studies and a Virtual Biorepository

Unfortunately, we can't immediately start collecting every type of data we'd like SEER-wide. That is why we have launched a series of pilot studies that can help us better understand the barriers

and challenges to collecting these new types of data.

Take, for example, the aforementioned supercomputers. They are part of a pilot effort being conducted in collaboration with the US Department of Energy (DOE).

Under this partnership, we are working with DOE scientists to develop tools that will allow SEER to collect the data elements it has traditionally captured (e.g., cancer type and grade) as well as new data elements (e.g., biomarkers, recurrences) from patient medical records. Compared with the manual collection that is often used now, automated data collection would be far more reliable and rapid, even if some cases still require manual coding.

This pilot with the DOE will allow us to assess the feasibility of this automated data collection and refine these new data collection tools, ensuring, for example, that they are accurately capturing the data we're looking for.

Another SEER pilot study is focusing on collecting data on the use of oral cancer drugs. Such data have traditionally been difficult to amass because these treatments are not given at hospitals or at doctors' offices but are obtained through pharmacies and taken by patients at home. Little is known about the use of these drugs and adherence to prescribed regimens in a real-world population—that is, outside of a clinical trial setting.

In one such pilot study being conducted with the SEER registry in Georgia, linkages have been established between the registry and all Walgreens and CVS pharmacies in the state. These linkages connect a patient's information in the pharmacy's records to the same patient in the SEER registry, allowing for the real-time collection of data on filled prescriptions for oral cancer therapies.

Having these types of real-world data readily available would be invaluable. From them, researchers could glean important insights into usage patterns that, for example, can identify whether certain subgroups of patients are not getting the therapies they need.

Similar pilot efforts are creating linkages between SEER registries and health insurance claims data from some of the largest health insurance companies in the country. These linkages are similar to those that enable the SEER–Medicare Linked Database, which was established in the early 1990s. This database has been the primary data source for more than 1,600 published studies, including those that have provided information on important new cancer incidence trends and long-term survivorship issues.

Creating these linkages with commercial health insurance claims will allow SEER to capture much more information on people diagnosed with cancer who are under the age of 65 (i.e., not yet in Medicare), including critical data on treatments that patients receive over time. Such information can offer a window into real-world clinical practice and its potential implications for patients and the health care system at large.

Other pilot linkage efforts underway involve collecting data on the results of genomic tests,

including multigene panels like the Oncotype DX test used in the TAILORx breast cancer trial. As oncology continues on the road to precision medicine, the ability to systematically collect these data would go a long way toward informing this movement.

Although these linkage pilots are noteworthy, another exciting enhancement to SEER under active evaluation is on the other end of the research spectrum: the development of a "virtual biorepository."

The aim of this repository is to provide information on tumor samples stored at institutions across the country. This will allow investigators to search for samples from patients with certain demographic or clinical characteristics or certain outcomes, and then request those samples (including all related clinical data on that patient) via an honest broker for use in their research.

Currently, six SEER registries are participating in a pilot study of the virtual biorepository, focusing on specific survivor groups for breast and pancreatic cancer. The lessons learned from this pilot will inform our efforts to eventually scale up to a larger virtual biorepository.

New Data, Important Opportunities

The enhancements planned for SEER are exciting and much needed.

Of course, patience will be an important ally moving forward. It will take time to implement these ambitious plans, and there is always the possibility that—given the scale of these efforts—they won't always go according to plan. In fact, performing these activities using small pilots allows us to understand where the barriers and challenges might be and to realistically assess whether they can move past the pilot phase.

That said, we are taking the necessary steps to eventually make these new tools and resources integral parts of SEER. In so doing, we will be opening up new doors to cancer researchers and greatly expanding the value of this unique and needed program.

Over the coming decade, stay tuned to hear more about a remarkable array of new scientific opportunities. As always, we welcome ideas and feedback from the cancer community about ways to strengthen the utility and usability of SEER data for both cancer research and cancer control planning.

This post was originally published on August 23, 2018, by the National Cancer Institute. It is republished with permission.

© 2026 Smart + Strong All Rights Reserved.
http://beta.docker.cancerhealth.com/blog/enhancements-ncis-seer-program-creating-new-research-opportunities