

2022 NPCR OKLAHOMA SUCCESS STORY

Oklahoma Central Cancer Registry; Christy Dabbs, AA, CTR; Spencer Denison, MPH; Raffaella Espinoza, MPH

Using Hospital Discharge Data to Determine Accurate 2020 Cancer Incidence During COVID Pandemic and Improve Reporting for Oklahoma

National Program of Cancer Registries SUCCESS STORY

SUMMARY

In 2020, the cancer screening dropped dramatically nationwide due to the COVID pandemic¹. It is thought that due to the drop in screenings, medical staff shortage, public fear of going to hospitals for non-emergency procedures, etc., that there would be a drop in cancer diagnosis as well. This, along with the loss of a registry team member, the Oklahoma Central Cancer Registry (OCCR) decided to rely on Oklahoma hospital discharge data (HDD) to assess for 2020 compliance and assist Oklahoma reporting facilities with cancer case finding and reporting. The OCCR linked 2020 HDD, collected by the Oklahoma State Department of Health, with the cancer registry database for diagnosis year 2020. Spreadsheets were distributed to the respective reporting hospital for review then returned to the OCCR. Reporting hospitals indicated missed cases that would be abstracted and submitted. The linkage was challenging but additional missed cases were reported to OCCR.

CHALLENGE

The OCCR uses the average number of cases reported within the last three years to determine the expected number of cancer cases to be reported by facility. A facility is considered compliant with reporting if at least 95% of the expected number of cancer cases are reported. The Compliance and Education Specialist will assess for compliance quarterly and send out letters to the facilities on their reporting progress. However, the diagnosis year 2020 presented many challenges which possibly impacted the diagnosis of new cancer cases, as well as healthcare facilities ability to catch up with reporting due to critical staff shortages². In addition, the Compliance and Education Specialist resigned which proved challenging to the OCCR in assessing for compliance.

The data linkage between HDD (Inpatient discharges from hospitals, outpatient surgery discharges from hospitals, outpatient surgery discharges from free-standing ambulatory surgery centers) and the cancer registry database posed some additional challenges:

1. De-Duplicating Hospital Discharge Data:
 - HDD can contain multiple encounters from the same hospital for the same patient.
 - HDD uses ICD-10-CM codes whereas cancer registry uses ICD-O-3 topography codes.
2. Linking HDD with the Registry Database:
 - Hospital names on the HDD do not necessarily correspond exactly with hospital names in the cancer registry database.
 - ICD-10-CM codes do not always translate exactly to ICD-O codes.
 - SAS 9.4 versus Match*Pro for linkage.

SOLUTION

The OCCR decided to rely on Oklahoma HDD in assisting Oklahoma health care facilities with the reporting of 2020 diagnosed cases, as well as assuring completeness of cancer data. The linkage process was adjusted as challenges arose:

1. The HDD was first filtered to exclude any non-reportable cancer diagnosis. The reportable cancer cases were taken from the National Cancer Institute - SEER Case finding List³. Deduplication of the data was conducted through SAS which would exclude any encounter in which an individual was diagnosed with the same primary cancer diagnosis at the same facility. However, if an individual was diagnosed with the same cancer at multiple facilities, then each encounter would be retained. Once the data was deduplicated we used the North American Association of Central Cancer Registries (NAACCR) Death Clearance Manual Appendix J to convert the ICD-10-CM codes to ICD-O-3 codes using SAS.

2. We quickly realized that SAS was not the optimal program to perform the linkage, and decided to use Match*Pro. Several attempts were made to utilize SAS to link the HDD file to the registry database file. To use SAS' linkage capabilities, the variables from the two different files would have to be uniform in format and content. Both files would have matching variables, like social security numbers, date of birth, last names, first names, middle initials, facility of diagnosis, and cancer diagnosis that would give high confidence in a positive match. It was because of their differences in source uploads and system standardizations that would be additional time and resource to overcome, i.e., HDD SSN length ranged from 1-7 digits while the cancer registry data was set at 7 digits. Additionally, the SAS linkage would be a deterministic link, which produces large number of false negatives when data items are missing or have incorrectly typed in info. Match*Pro uses probabilistic record linkage approach which outputs the percentage certainty of matches based on inputted linkage criteria and variables. By using this program, it provided the flexibility needed for the differences between data items like social security number, as well as accounting for minor typos in data content.
3. Hospital names on the HDD spreadsheet were manually translated to the matching hospital name in the registry database to facilitate a more accurate linkage in Match*Pro.

RESULTS

The deduplicated HDD file contained 25,091 cancer related encounters and was linked to 20,749 cancer cases from the registry database. The initial linkage resulted in 10,467 matches and 2,484 uncertain matches and the remainder were non-matches. Due to time and resources, not all uncertain or non-matches were reviewed, therefore staff concentrated

on small volume hospitals. The review of the completed spreadsheets was conducted and combined on a master spreadsheet for tracking and quantification of cases already reported, missed cases and non-reportable cases. It is possible to link the HDD with cancer data to find missed cases from reporting facilities. Additional time was spent in preparing the files prior to linkage, reviewing the linkage results and preparing the spreadsheets to be distributed to the reporting hospital. However, the link resulted in capturing missed cases, and the ability to assess for data reporting and completeness more accurately during a worldwide pandemic.

SUSTAINING SUCCESS

New processes can often require extra time and resources, especially during the pilot phase. However, the OCCR plans to refine and automate as much of the deduplication and linkage process as possible, as well as revisit the Match*Pro configuration file. The goal is to conduct the HDD/cancer data linkage quarterly and provide these spreadsheets to Oklahoma healthcare reporting facilities in hopes of improving timeliness and completeness in cancer case reporting to the OCCR.

STORY QUOTE

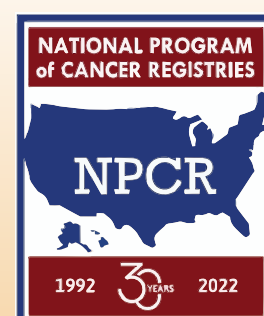
"High quality and complete data are not an accident, but the result of intention, effort, and skillful execution. Good data leads to better decision-making for cancer registries and ultimately for data users." – Raffaella Espinoza, Interim Program Director, and Spencer Denison, Epi Support

REGISTRY CONTACT INFORMATION

405-426-8030
Oklahoma Central Cancer Registry Website

REFERENCES

1. National Cancer Institute. *Working to Close the Cancer Screening Gap Caused by COVID*. 17 May 2022, <https://www.cancer.gov/news-events/cancer-currents-blog/2022/covid-increasing-cancer-screening>
2. Impact of the COVID-19 pandemic on the hospital and outpatient clinician workforce: challenges and policy responses (Issue Brief No. HP-2022-13). Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. May 2022.
3. National Cancer Institute- Surveillance, Epidemiology and End Results. *Casefinding Lists*. <https://seer.cancer.gov/tools/casefinding/>



U.S. Department of
Health and Human Services
Centers for Disease
Control and Prevention