To leverage existing Surveillance Epidemiology and End Results (SEER) data elements to investigate the development of an algorithm using data items collected and transmitted through SEER to determine the likelihood that a patient received neoadjuvant treatment.

Neoadjuvant therapy, also referred to as induction therapy, is generally defined as systemic therapy given before localized cancer treatment. The use of this treatment is increasing and it remains underutilized in some populations.1-5 There currently is no formal collection of this type of treatment data—largely due to the fact that a standardized definition for neoadjuvant data collection does not exist in the literature.

Although treatment sequence variables are collected, the quality of these variables is under review by NCI’s Data Quality Analytics and Interpretation Branch. Routine and accurate collection of this treatment sequence is essential to better understand therapeutic effectiveness and guide strategies in treatment plan for cancer care.

Leveraging Current SEER Data Elements to Characterize Receipt of Neoadjuvant Treatment
Melissa Bruno, MPH, Clara Lam, PhD, Serban Negoita, MD, DrPH

Aims

To leverage existing Surveillance Epidemiology and End Results (SEER) data elements to investigate the development of an algorithm using data items collected and transmitted through SEER to determine the likelihood that a patient received neoadjuvant treatment.

Background

Neoadjuvant therapy, also referred to as induction therapy, is generally defined as systemic therapy given before localized cancer treatment. The use of this treatment is increasing and it remains underutilized in some populations.1-5 There currently is no formal collection of this type of treatment data—largely due to the fact that a standardized definition for neoadjuvant data collection does not exist in the literature.

Methods

- SEER 2010-2016 colon cancer cases
- Dataset= SEER 20 Regs, Nov 2017 Sub (1973-2016)
- Sex= Male, Female (exclude unknown)
- Age= All ages (exclude unknown)
- Year of diagnosis= 2010-2016
- Registries= All registries
- CS Schema v0204 = colon
- 18 NAACCR vol. II variables reported and submitted to NCI
- These variables were re-coded into neoadjuvant categories based on their indication of the likelihood of receiving neoadjuvant therapy.

Figure 1. Example of the re-coded NAACCR Chemotherapy variable into the neoadjuvant categorizations for use in an algorithm.

Figure 2. Breakdown of neoadjuvant coding categories for the neoadjuvant algorithm variables.

Table 1. Breakdown of neoadjuvant coding categories for the NAACCR treatment sequence variables.

Table 2. Neoadjuvant algorithm performance compared to NAACCR sequence variables. Codes 2-9 cases cover possible, likely, unindicative, & unknown neoadjuvant categories.

Conclusions & Next Steps

Following these preliminary algorithm results, the following next steps are planned for its improvement and testing:

- Adding additional elements to the algorithm calculations, for example, the collaborative stage evaluation variables.
- Testing the algorithm on other cancer sites, for example, breast.
- Validating the algorithm’s performance using the SEER*Medicare linked dataset.

References