

Clinical Data Partner Activity

Kayla Xu, Joshua Gabbay, Avinash Chauhan

1. Categorical variables are data that can be sorted into different categories (color of flower). Discrete variables are data that are countable in integers (number of cars). Continuous variables are data that can be any real number (height).
2. Breast_carcinoma_primary_surgical_procedure_name
3. Depends on the initial surgery you have (lumpectomy, mastectomy, other). Categorical variable.
4.
 - a. <https://www.nature.com/articles/s41523-022-00456-4>
 - i. Mastectomies are correlated with more pain and higher rates of inflammation in patients than lumpectomies do.
 - b. <https://link.springer.com/article/10.1245/s10434-021-10499-9>
 - i. This study reported differential outcomes between two-dimensional and full three dimensional specimen imaging in evaluating lumpectomy. In short, the three-dimensional volumetric specimen imager demonstrated an improved correlation between the main lumpectomy specimen margin status and surgical pathology.
5. Anatomic treatment site – where breast cancer is located within the breast.
6.
 - a. The different surgeries will be more prevalent depending on the anatomic treatment site of radiation.
 - b. Mastectomies are associated with a lower survival rate of breast cancer.
 - c. Distant recurrence and distant site are more associated with lower survival rates of breast cancer.
7. From plotting the two variables against each other, we saw that most patients get radiation treatment at the primary tumor field, with most of those patients getting a lumpectomy before. With patients receiving radiation treatment at the regional site, most of those patients are likely to get a modified radical mastectomy. The survival analysis with breast carcinoma surgical procedure showed that patients that do not get surgery have the lowest survival rates, and patients who got a simple mastectomy had the highest survival rate. The survival analysis with anatomic treatment site showed that patients undergoing treatment at a local recurrence site have the lowest survival rate, while patients undergoing treatment at a regional site or the primary tumor field have the highest survival rate. The distant site doesn't seem to have enough data to say anything conclusive about the survival rate.