

Developing Tools to Improve Data Quality

AUTHORS: Verónica Rivera-López, ODS, and John Orr

SUMMARY

The Virgin Islands Central Cancer Registry (VICCR) has been collecting cancer data for Virgin Islands (VI) residents since 2016. VICCR still performs all of the territory's abstracting. As a group with very limited human resources, collection of quality data has been challenging. This year we created a tool to help improve data quality.

CHALLENGES

VICCR had many cases where the EDITS would pass in an abstract, but when reviewing the case, information is missing or not consistent. Examples of missing or incorrect information include:

- Sequence number for patients with multiple tumors.
- Incomplete date of diagnosis (specifically day).
- Missing or incomplete address (county and estate, which is a custom data item).
- Unknown race.
- Vital status (alive) in patients older than 90.

SOLUTION

The Quality Tool identifies missing or incomplete items and produces reports that are used to improve the data quality. It can:

- Generate productivity reports that track abstracts by year, month, and abstractors. The report identifies areas for improvement.
- Evaluate monthly diagnoses by year, month, and provider. The report helps identify possible missing cases, allowing the VICCR to follow back with the providers (mainly physician offices).
- Check for possible duplicates in the database, which is a process that we perform during the year, not just for data submission.
- Check for missing birthplace to reduce the cases with unknown race by comparing the report with the hospital's electronic health record. By knowing the birthplace and following the SEER Program Coding and Staging Manual – Appendix D: Race and Nationality Descriptions, race can be recoded.
- Checks for tumor behavior to keep track of benign, *in situ*, and invasive tumors. This report was created in response to a reported error in the data submission file.

RESULTS

- The number of cases with unknown race has been reduced to within the required limit.
- The amount of unknown information in the abstracts has been reduced.
- Data preparation time at submission is reduced due to tracking of missing information throughout the year.

CONCLUDING REMARKS

Although the tool is still in development, it has helped improve data completeness and quality.