

Evaluation of LexisNexis Batch Solutions for Cancer Registries in New York State

Eva Pradhan, Francis P. Boscoe
New York State Cancer Registry

Introduction

Background: Incomplete or inaccurate demographic information can impact the results of epidemiological analyses. Despite the best efforts of registry staff, a small share of the demographic information for cancer patients can never be determined with certainty. While this share has decreased over time because of improved information technology, some of the gains have been offset by increasing reports from non-traditional sources such as laboratories and private physicians.

It can take a tremendous amount of time and resources for cancer registries to individually research these cases to get the correct information. LexisNexis, an Ohio-based division of the Anglo-Dutch publishing company Reed Elsevier, offers batch solutions to determine this information and returns results in a fraction of the time required for manual review, at a cost of just pennies per record.

Purpose: To assess LexisNexis Batch Solutions as a resource for determining the date of birth, address at diagnosis, and/or social security number of a sample of patients reported to the New York State Cancer Registry (NYSCR).

Methods: We generated the following files:

- (1) addresses missing census tract for cases diagnosed between 1997-2010 (n=23346, 1.5% of total tumors diagnosed in that period)

Results

- (2) partial birthdates and ambiguous birthdates (2 or more sources with birthdates differing by >1 year) diagnosed between 1976 and 2012 (n=8800, 0.3% of all cases diagnosed in that period).
- (3) cases diagnosed between 2003-2010 with more than one non-blank social security number (n=5958, 0.7% of all cases diagnosed in that period).

The time periods were chosen based on current registry priorities and the available budget.

Results:

For the address searches, 10,133 (43.4%) had address information corresponding to the time of diagnosis. Of these, 4332 (42.8%) were geocodable addresses (containing a house number and street) and were used to update information on our registry. Additionally, 487 non-geocodable addresses provided by LN were used in combination with existing registry value. For instance, when our registry had PO Box information and LN provided a street name, we used the combination of those addresses. Most of the address searches were for older cases, with 92% from 2003 or before.

For the birthdate searches, LN provided results for 3237 (36.8%) of the cases. We were able to use 3161 (98%) of these. Most of those birthdates (n=2967) were complete and used as provided by LN. Some of the partial birthdates (n=194) were useful to resolve ambiguous existing values in the registry.

For the SSN searches, 5499 (92.3%) were found. We were able to use 100% of the SSNs provided by LN to verify or replace existing values.

Fig 1. Usefulness of LexisNexis results

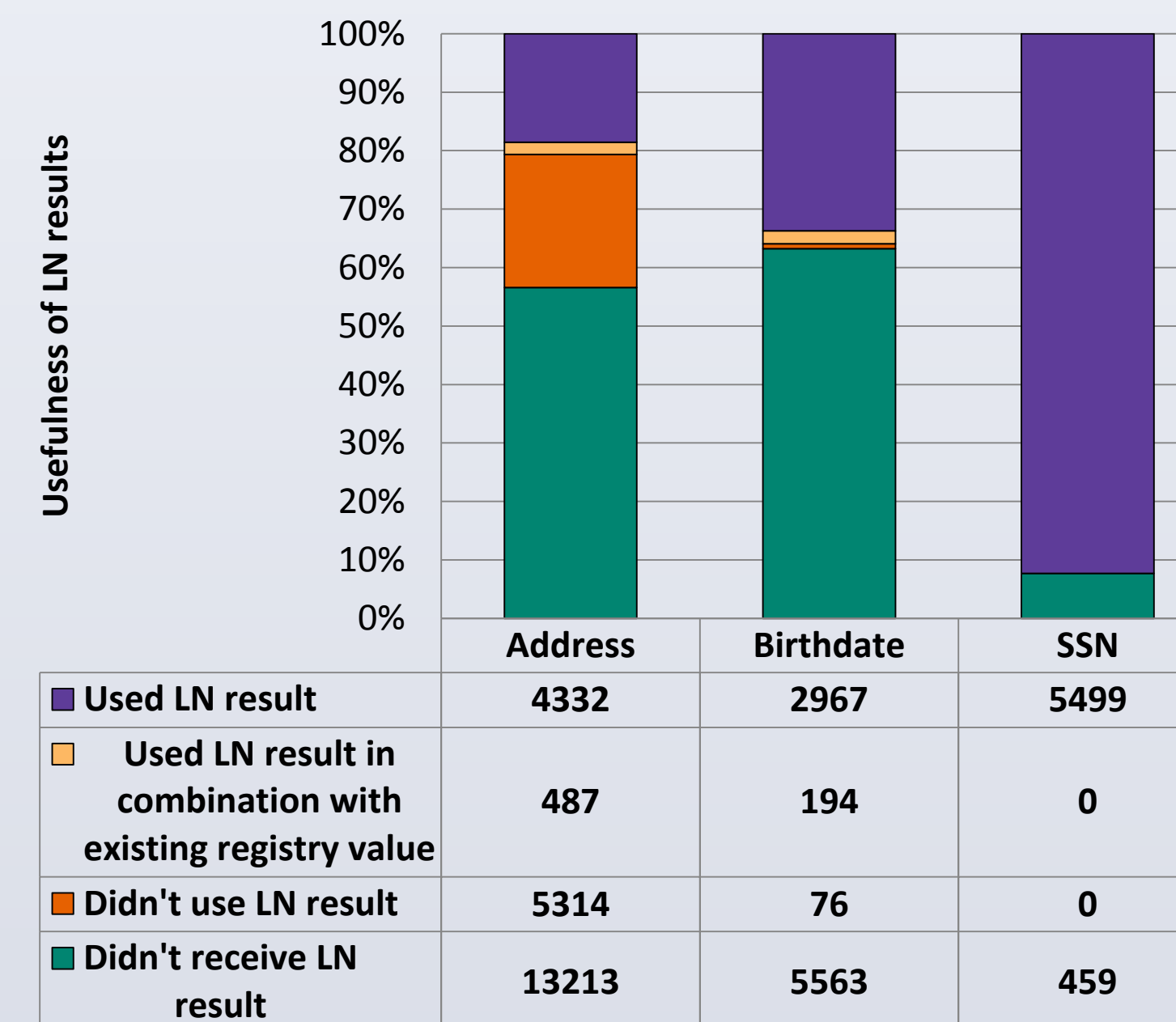
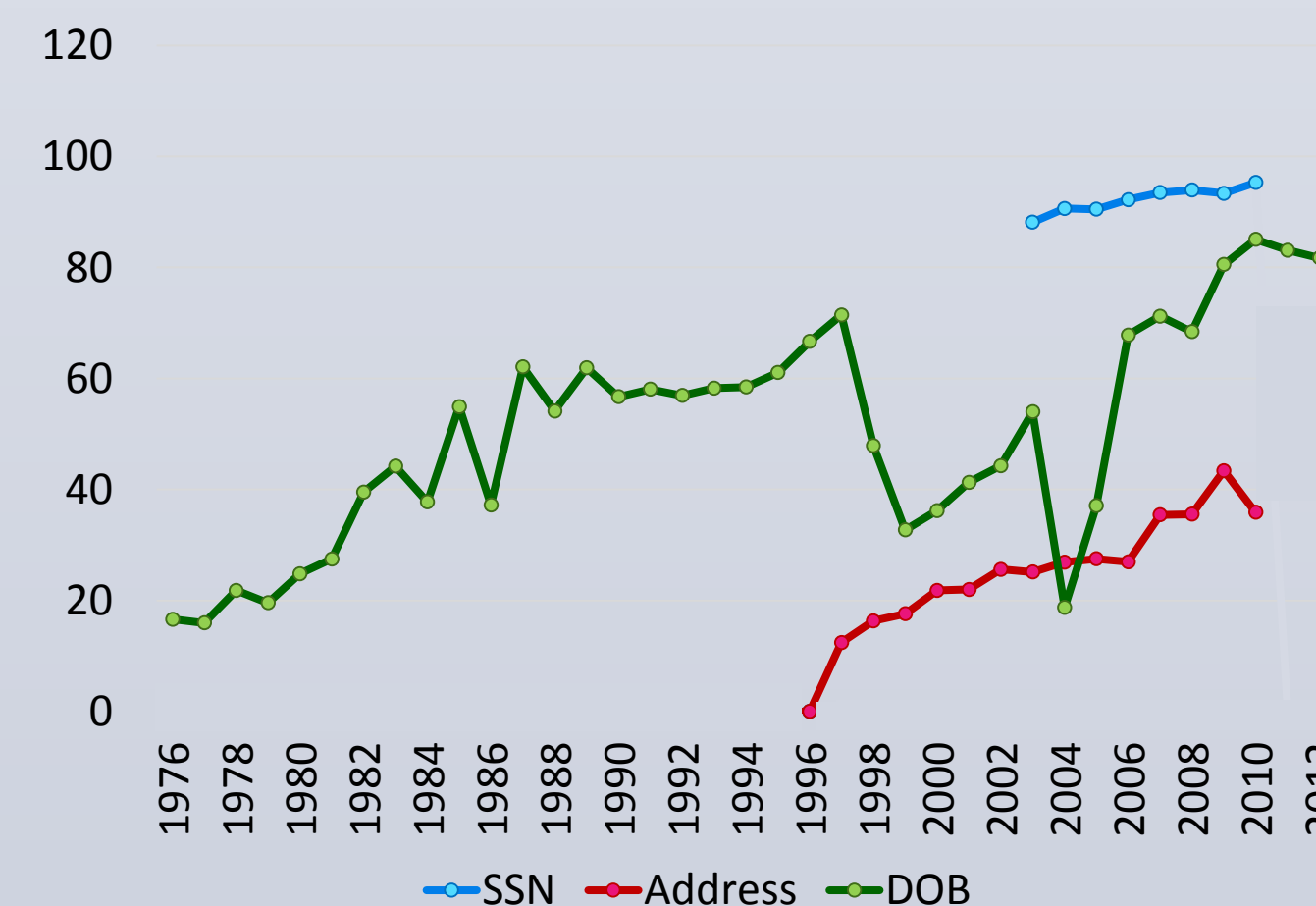


Fig 2. Success rate of LN, by diagnosis years



Discussion and Conclusion

Having complete and correct demographic and identifiable information is key to cancer registries, for the purpose of matching cases with other databases, for correctly summarizing population characteristics, and for maintaining data quality for cancer registries. LexisNexis offers a great cost-effective solution for obtaining such information with a quick turn-around time.

The success rate for this analysis was defined as the number of cases/tumors for which the values provided by Lexis was used either to replace the existing blank/non-blank value or to resolve ambiguous values. As seen in Fig 2., the success rate of LN slightly increased with the diagnosis years for all three variables, but even very old cases returned usable results.

The success rate indicated in this poster may underestimate the effectiveness of LN that would be experienced by other registries, as the cases used in this study represent the hardest cases which had already been worked on by registry staff members to try and resolve, in some instances going back decades.

Acknowledgement

The authors are thankful to the New York State Cancer Registry for providing opportunity to work with LexisNexis.

Contact information

Eva.pradhan@health.ny.gov
Francis.boscoe@health.ny.gov