Development of data linkage strategies to study early life exposures and breast cancer in young California women.

Reynolds P, Hurley S, Layefsky M, Quach T. Northern California Cancer Center, Fremont, CA.*

Background: Evidence is emerging for the importance of pre- and peri-natal environmental exposures in the etiology of adult-onset diseases, including breast cancer. Such factors, however, are extremely difficult to study given that such exposures are impossible, or nearly impossible to self-report. Consequently, the development of methods to link adult health outcome data, birth data, and historical socioeconomic (SES) and environmental contaminant data, could prove extremely useful in studying such exposures.
Objective: As part of a larger case-control study aimed at evaluating the association between breast cancer risk and early-life exposures, the purpose of this analysis is to evaluate the feasibility of linking female breast cancer cases to California birth records.
Methods: All cases of breast cancer diagnosed in California women, 1988-2003 who were born 1960-1969 were identified from the California Cancer Registry (CCR). A group of women with other cancer diagnoses (i.e., non-reproductive cancers and non-smoking-related cancers) were similarly selected to serve as controls. Linkage of female cancer cases to their birth records is complicated by name changes due to marriage and incomplete data on birthplace available in the CCR. The CCR has a maiden name available for only 15% of married women diagnosed with breast cancer. Birthplace is available for less than 60% of the cases. In order to maximize the success of our CCR linkage to birth records, we have gathered marriage data both from the state and from 11 of the 12 largest counties in California. These data are being linked to the CCR data to harvest additional maiden name and birthplace information in order to improve the success of our linkage to birth records, thereby increasing our study's sample size.
Results: Linkages of the CCR to the marriage and birth data are currently underway. Initial linkage of cases listed as never married or married with a maiden name available in the CCR resulted in an overall match rate of approximately 41%, with success considerably higher (63%) among those records with a California birthplace listed in the CCR. After augmenting CCR records with additional information on maiden name and birthplace from linkages to the marriage files, we have increased the number of cases linked to a birth record (and therefore eligible for our case-control study) by more than 50%.
Conclusions: Preliminary results suggest that augmenting CCR data with county and statewide marriage data is useful for improving our ability to link female cancer cases to California birth records, thereby enhancing our ability to study early life exposures in breast cancer etiology.