Introduction:
1. Relative survival methodology with presumed-alive assumption has been adopted by NPCR and NAACCR for population-based cancer survival analysis. However, Pinero et al. (2014) and Johnson et al. (2010) have observed overestimation biases with use of these methods.
2. The survival time of a cancer patient is directly determined by date of diagnosis and date of last contact. Any reporting issues in these variables, such as incidence reporting delay, may cause biases in survival outcomes.
3. The purpose of this study is to assess the impact of incidence reporting delay to identify and quantify potential biases in population-based survival analysis with presumed-alive assumption.

Study Data:
1. Incidence reporting delay was calculated from NPCR data submissions 2001-2015 from 16 registries who met NPCR data quality standards, demonstrating consistent patient IDs between submissions, and performed NDI linkages.
2. The foundation population included the malignant cases from the 16 registries in the NPCR 2015 data submission with age between 0-99.
3. Two study populations were formed from the foundation population. a) Population 1: Cases diagnosed between 2000-2006 with follow-up to 2006; b) Population 2: Cases diagnosed between 2001-2007 with follow up to 2007.

Method:
1. To isolate the effect of incidence reporting delay, the study populations used values of sex, race, date of diagnosis, date of last contact, vital status, age of diagnosis from the foundation population throughout the study.
2. Population 1 and Population 2 were analyzed through the initial cohorts of Population 1 and 2 are as color coded. The cases from submission 2010-2015 are the reporting delayed cases sequentially added to the starting cohorts.

Results:
Table 1: Incidence reporting delays between 1998 and 2007 in NPCR data submission 2001-2015 from 16 registries. The number at the top of each column represents the cases reported without delay in a 24-month reporting window of a diagnosis year. The initial cohorts of Population 1 and 2 are as color coded. The cases from submission 2010-2015 are the reporting delayed cases sequentially added to the starting cohorts.

Table 2: The effects of incidence reporting delay on 5-year relative survival outcomes of Population 1 are demonstrated with colon & rectum and leukemia. The underestimation of survival is the net difference between survival estimates of earlier submission years and Submission 2015. The reporting delay has no effect on colon & rectum, but cause significant underestimation of survivals in leukemia. Leukemia survival estimates stabilized after 4 years of subsequent submissions.

Table 3: The effects of incidence reporting delay on 5-year relative survival outcomes of Population 2 are demonstrated with colon & rectum and leukemia. The underestimation of survival is the net difference between survival estimates of earlier submission years and Submission 2015. The reporting delay has no effect on colon & rectum, but cause significant underestimation of survivals in leukemia.

Table 4: Six cancer sites whose survival estimates are significantly affected by the incidence reporting delay in Population 1 and 2. Population 1 included diagnosis year 2000 which has more reporting delay than later diagnosis years, tends to have slightly greater biases than population 2. These sites need additional submissions to achieve stabilized survival estimates in both study populations, especially all-sites-combined, myeloma, and leukemia.

Conclusions:
1. For a mature population-based cancer surveillance system, assuming incidence reporting delay is the single source of bias, survival is under-estimated with the presumed-alive assumption.
2. Recent NPCR data have less, yet more stable, incidence reporting delay. Using diagnosis periods that avoid early NPCR data in the study population may reduce the underestimation bias; and may also significantly increase the survival estimates for cancer sites experiencing severe incidence reporting delay.
3. The current NPCR surveillance database included diagnosis year 2001 to the most recent diagnosis year - 1 and corrects 3 less impacted cancer sites. However, for all-sites-combined, myeloma, and leukemia, the period is not enough to mitigate their underestimation. Because of the short history of NPCR data, the study can’t evaluate more recent diagnosis years. It would be beneficial to conduct an annual assessment of such biases with each new data submission.
4. It is necessary to continue to identify and quantify other sources of biases that may contribute to the overestimation biases reported in the literature.

Contact:
Xing Dong: xdong@icf.com
Reda Wilson: dfow@cdc.gov

Impact of Cancer Incidence Reporting Delay on Population-Based Survival Analysis
X Dong*, Y Ren, K Zhang, R Wilson
1Division of Cancer Prevention and Control, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, GA
2ICF International, Fairfax, VA

National Center for Chronic Disease Prevention and Control and Health Promotion Division of Cancer Prevention and Control