**Introduction**

1. National Program of Cancer Registries (NPCR) of CDC funds 48 central cancer registries in the U.S. and its territories.
2. Historically, the cancer incidences reported in the NPCR are not adjusted for reporting delay and reporting errors.
3. Reporting delay happens when a case reports to CDC after its designated 24-month reporting window. Reporting error is removing a case from a subgroup population on a set of predefined criteria.
4. The NPCR reporting delay model is to explore a better modeling approach to estimate reporting delay distributions that suits the characteristics of the NPCR data.
5. The resulting predicted reporting delay distributions can be used to accurately describe the cancer incidences in the NPCR data.

**Methods**

- **Data:**
  1. Reporting delays and errors defined similarly to what Midthune et al. described in 2005.
  2. The NPCR data used in the model development included 23 central registries and submission 2000 to 2010.
  3. Variables used to define reporting delays and errors were registry, primary cancer site, sex (male and female), race (white, black and other), age group (18 groups, 0-4, 5-9, ..., 85+), and year of diagnosis (1998 to 2007).
  4. The NPCR model was initially developed with melanoma (white only) incidences in the NPCR data. It was later applied to all other cancer sites. Additional sites reported here are prostate (male only), breast (female only), lung and colon & rectum.

**Table 1:** The reporting delays and errors patterns of melanoma in 23 central registries in the NPCR data. Note: Table 3 structured similarly to the table published by Midthune et al. in 2005. The net total is the accumulated melanoma incidences by submission 2010 (Net total = Initial + (Total Delay) – (Total Error)).

- **Methodology:** The NPCR net truncated delay model (fixed effect model)
  1. The NPCR delay model assumed that the net differences between reporting delays and errors are independent Poisson distributions.
  2. The net differences can also be considered as binomial distributions conditional on the total reported incidences.
  3. The conditional reporting delay probabilities can be estimated with the generalized linear model under the binomial distribution assumption with complementary-log-log link function.
  4. Predictor: indicators of delay time period, such as 2, 3, ..., or 11 years of delay in reporting.
  5. Variances of predicted delay distributions can be estimated as Brookmeyer et al. proposed in 1990 with Greenwood’s formula.

**Results and Discussions**

**Table 2:** Comparisons of the predicted delay distributions between the NPCR fixed effect delay model and the NCI fixed effect delay model of 5 cancer sites in submission 2010. The predicted reporting delay distributions between the NPCR and NCI models are similar to each other in 5 cancer sites tested. There is resemblance between the predicted delay distributions of melanoma and prostate cancer. It is noticeable that breast cancer has less reporting delays than prostate cancer in the NPCR data. The upward reporting delay adjustment to the incidences diagnosed in 2007 and reported in 2010 ranges from 4.7% to 10.8% in 5 cancer sites tested.

**Table 3:** Comparisons of the effects of delay adjustment on melanoma incidence rates between the NPCR fixed effect delay model and the NCI fixed effect delay model. The effects of reporting delay adjustment can be seen more obviously in Figure 1.

**Figure 1:** Estimated incidence rates from Sancer sites using the NPCR fixed effect delay model. (A) Estimated prostate cancer incidence rates. (B) Estimated breast cancer incidence rates. (C) Estimated lung cancer incidence rates. (D) Estimated colon and rectum cancer incidence rates. (E) Estimated melanoma incidence rates.

**Conclusions**

1. The NPCR fixed effect delay model is extended from the NCI fixed effect delay model with simpler model specifications. The estimation process is stable and less prone to non-convergence.
2. The NPCR fixed effect delay model and the NCI fixed effect delay model agreed with each other in the predicted reporting delay distributions and the corresponding adjustments to the cancer incidence rates.
3. The NPCR fixed effect delay model is a good alternative approach to model reporting delay at the national level in the NPCR data.
4. The NPCR fixed effect delay model can be applied to states if the data quality at the state level meet the requirements for modeling.

**References**


**Acknowledgement**

Special thanks to:
1. Douglas Midthune, Mathematical statistician at NCI
2. Dr. Baoguang Song, Mathematical statistician at CDC/NCHSTP/DPAP/QSDMB