1979 and closing. After the 2005 update, 1,052 subjects with vital status previously unknown were found to have survived into 2002 or beyond. No NDI matches were identified among these 1,052 subjects giving a match specificity of 100 percent. A full breakdown of the cohort by January 1, 2002 vital status as assessed in June 2003 and again in April 2005, with ascertainment of deaths by the NDI and the NWTS is given in Figure 1. A further cross-tabulation of NDI match status by January 1, 2002 vital status is shown in Table 1. The individual with unknown vital status whose record was not sent to the NDI had a recorded date last seen that changed from after January 1, 2002 to before that date during the mortality update.

Table 2 shows match rates for the 789 known decedents. Using multiple logistic regression, three factors were independently associated with a lowered match rate: lack of SSN, Hispanic ethnicity and birth outside the US. Sensitivity was 96.7 percent for subjects with an SSN on file but only 87.8 percent for those without. Sensitivities were also substantially lower for Hispanics (76.4 percent) and for those born abroad (56.5 percent) compared to those of non-Hispanic Caucasians and US born, respectively. Subjects who died during earlier calendar periods were also less likely to have a match, but year of death was not a statistically significant factor after adjustment for the confounding effects of SSN. Age at death was not significantly associated with the probability of having a match. This was true when age at death was included as a continuous variable ($p = 0.16$) or as a grouped linear variable ($p = 0.08$).

Results of the survival analyses comparing standard NWTS, NDI supplemented and NDI only follow-up are presented in Table 3. These are presented separately for the "early" (NWTS-1,2) and "modern" (NWTS-3,4) treatment eras. After augmenting the NWTS data with NDI search results, 2,351 subjects with vital status unknown on closing had an average of 5.6 years of follow-up added to their records between 1979 and 2001 for a total of 13,166 person-years. Eighteen deaths based on the NDI match were added to the 198 pre- and 789 post-1979 deaths ascertained by the NWTS. The NDI augmentation led to small increases in each of the actuarial survival estimates.

Since the actuarial procedure for the augmented data increased the date last seen to January 1, 2002 for all surviving subjects, but increased the numbers of deaths only by those that occurred after 1978, part of the apparent increase in survival was due to failure to account for deaths before 1979 in subjects lost to follow-up earlier. However, only 43 person-years of follow-up, during which 2.3 deaths would have been expected, were added to the pre-1979 period as a result of the NDI search. Consequently, any bias caused by the gap before the NDI became operational would be minimal.

The NDI only analysis removed 80 deaths, added in a total of 1,170 person-years of follow-up for these 80 subjects and resulted in additional increases in the survival estimates. The improvement in survival from the first two NWTS trials to the second two is apparent no matter which death ascertainment method is used.

Among the 2,351 subjects whose follow-up was augmented by NDI search data we would have expected 79.3 deaths between 1979 and 2001 based on NWTS mortality rates. This is 4.4 fold higher than the 18 additional deaths actually found. SSNs were available for 23 percent of known decedents and 58 percent of subjects with an unknown vital status. The NDI search identified 6/12.1 (50 percent) of the deaths expected among participants with an SSN but only 12/67.3 (18 percent) of the deaths expected among those without. The difference between observed and expected deaths was not statistically significant among participants with an SSN ($p = 0.09$) but was among those without an SSN ($p < 0.001$).

Participants of non-Hispanic Caucasian ethnicity and those born in the United States had much better follow-up rates than children of other ethnicities and places of birth. By 10 years after enrollment 9.5 percent [95% CI: (8.5, 10.7)] of non-Hispanic Caucasian participants were lost to follow-up while 22.4 percent (19.5, 25.8) of African American, 19.5 percent (10.2, 35.4) of Asian and 25.9 percent (21.2, 31.4) of Hispanic participants had been lost. Of participants born in the US, 12.2 percent (11.1, 13.3) were lost by 10 years as compared to 35.5 percent (24.8, 49.2) of foreign born participants. No significant difference was seen across gender. Only 148 (6.1 percent) of the 2417 children with an SSN on file at the time of the NDI submission were lost to follow-up before 10 years, whereas 362 (17.2 percent) of the 2,105 without an SSN were lost. Since the date of SSN acquisition was not recorded, however, this precluded proper assessment of its effect as a time-dependent covariate on reducing loss to follow-up.

Conclusion

Three factors were independently associated with a lowered NDI match rate among known decedents: lack of SSN, Hispanic ethnicity and birth outside the US. Using NWTS mortality rates we estimated that 79.3 deaths should have occurred among subjects lost to follow-up yet only 18 such deaths were ascertained by the NDI search. We conclude that, in populations such as ours that are actively followed, the NDI cannot be used to reliably fill in missing follow-up data and that doing so may lead to inflated survival estimates.