

Validation of the Social Security Death Index (SSDI): An Important Readily-Available Outcomes Database for Researchers

James Quinn, MD, MS*
Nathan Kramer, BS†
Daniel McDermott, MD†

* Division of Emergency Medicine, Stanford University
† Division of Emergency Medicine, University of California, San Francisco

Submission history: Submitted May 29, 2007; Accepted September 19, 2007
Reprints available through open access at www.westjem.org

Study Objective: To determine the accuracy of the online Social Security Death Index (SSDI) for determining death outcomes.

Methods: We selected 30 patients who were determined to be dead and 90 patients thought to be alive after an ED visit as determined by a web-based search of the SSDI. For those thought to be dead we requested death certificates. We then had a research coordinator blinded to the results of the SSDI search, complete direct follow-up by contacting the patients, family or primary care physicians to determine vital status. To determine the sensitivity and specificity of the SSDI for death at six months in this cohort, we used direct follow-up as the criterion reference and calculated 95% confidence intervals.

Results: Direct follow-up was completed for 90% (108 of 120) of the patients. For those patients 20 were determined to be dead and 88 alive. The dead were more likely to be male (57%) and older [(mean age 83.9 (95% CI 79.1 – 88.7) vs. 60.9 (95% CI 56.4 – 65.4) for those alive]. The sensitivity of the SSDI for those with completed direct follow-up was 100% (95% CI 91 -100%) with specificity of 100% (95% CI 98-100%). Of the 12 patients who were not able to be contacted through direct follow-up, the SSDI indicated that 10 were dead and two were alive.

Conclusions: SSDI is an accurate measure of death outcomes and appears to have the advantage of finding deaths among patients lost to follow-up. [WestJEM. 2008;9:6-8.]

INTRODUCTION

Death has always been an important, if not the ultimate outcome in cohort studies, particularly for cardiovascular studies.^{1,2} Most of these have used direct follow-up of patients to determine death. This is not only a costly endeavor but often ends up with an inevitable number of patients lost to follow-up. Recently, federal databases have become available to investigators to track death as an outcome. The two readily available source databases for mortality are the National Death Index (NDI) and the Social Security Death Index (SSDI).

The advantage of these databases is that there is an incentive for the federal government to maintain their accuracy to help administer Social Security and other federal

benefits. As such, these databases are reported to be accurate within six months of a person's death, as it is mandatory for hospitals and funeral homes to report deaths so Social Security numbers (SSN) can be retired. The National Death Index (NDI) requires a formal submission and request for a third-party search at cost to the investigator. The SSDI has the advantage in that it is free and online.³ While others have compared the SSDI and the NDI databases as well as various search methods, the accuracy of the SSDI versus direct follow-up has never been assessed.^{4,5}

The purpose of this study was to validate the accuracy of the SSDI to determine whether it could be easily used as a source of death outcomes for researchers.

METHODS

After obtaining IRB approval, we identified a consecutive sample of patients presenting with syncope from an existing ED study database.⁶ Searching the SSDI, we took the first 30 patients who the SSDI listed as dead and the first 90 patients listed as alive after their ED presentation. The search was done at least one year after the initial ED visit to allow ample time for completion of the database. Although the database can be searched with incomplete information, for this study we used the patient’s complete SSN as well as full given name, surname and date of birth. If the patient was said to be dead, we requested the death certificate from the state in which they died (usually California) to confirm that it was a person from our study cohort. We then had a research coordinator blinded to the results of the SSDI search, complete direct follow-up. This was done by contacting the patients themselves, their families or their primary care physicians to determine their vital status and date of death. We used direct confirmation of death from these three sources as the criterion reference.

Statistical Analysis

We determined the sensitivity and specificity of the SSDI for determining death at six months, using direct follow-up as the criterion reference, and calculated 95% confidence intervals using 2x2 table and exact binomial distributions. The study was powered to have a 95% confidence interval width of 10% for the sensitivity.

RESULTS

All 30 patients deemed dead by the SSDI had confirmatory death certificates. There were a variety of causes of death in the selected cohort as illustrated in the Figure.

Of the 120 patients in the study we were able to contact 108 patient families or primary physicians to confirm vital status (90%). For the 108 patients with direct follow-up 20 were determined to be dead and 88 alive. Those dead at six months were more likely to be male (57%) and older [(mean age 83.9 years (95% CI 79.1 – 88.7) vs. 60.9 years (95% CI 56.4 – 65.4) for those alive)]. The sensitivity of the SSDI for those with completed direct follow-up was 100% (95% CI 91 -100%) with specificity of 100% (95% CI 98-100%) (Table). Of the 12 patients not able to be contacted through direct follow-up, the SSDI indicated that 10 were dead and two alive.

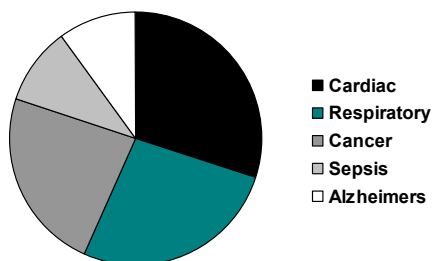


Figure. Causes of death reported on the 30 death certificates.

DISCUSSION

In this study we found that the readily available online SSDI was an accurate and facile database to determine death outcomes. It showed excellent sensitivity and specificity for those for whom we could complete direct follow-up, as well as information on patients who we could *not* contact, including 10 with confirmatory death certificates. If a study used mortality as a primary endpoint, 33.3% (95% CI 17% – 53%) of subjects who would ordinarily be reported as “lost to follow up” would be, in fact, positively identified by the SSDI as dead.

The database is available online and does not require cost or special expertise. However, its use does require institutional review board permission since protected health information (PHI) elements are used to search. We were able to secure a HIPAA (Health Insurance Portability and Accountability Act of 1996) waiver to complete the necessary follow-up for this and other studies.

The comparable federal database used successfully in previous research is the National Death Index (NDI). However, it takes a formal submission with a two-month response time and has associated costs. The fees for routine NDI searches consist of a \$350.00 service charge plus \$0.15 per user record for each year of death searched. For example, 1,000 records searched against 10 years would cost \$350 + (\$0.15 x 1,000 x 10) or \$1,850. A recent study showed the two databases to have comparable accuracy.² Other studies looking at the SSDI and NDI found results that were not as sensitive as ours. However, these studies compared databases and did not use direct follow-up as the criterion reference.^{4,5} One also used different search criteria.⁵

It is possible to search the SSDI database without a SSN. This can be done by name and other demographics. Studies not using the SSN in their searches found poorer sensitivity compared to when it was used, and investigators should beware of searches not using the SSN.⁴ Even compared to previous studies where the SSN was used, our sensitivity was better and may indicate that the maintenance and search tools of the database have improved over time.

Table. Performance of the SSDI to Determine Death Outcomes

SSDI	Dead	Alive	Lost To Follow-Up
Dead	20	0	10
Alive	0	88	2

Sensitivity 100% (95% CI 91% - 100%)
 Specificity 100% (95% CI 98% - 100%)

LIMITATIONS

For our syncope cohort, this was a secondary analysis carried out up to two years after the index ED visit. Hence, we were not able to determine how current the online SSDI is, nor verify the governmental claim that it is completed within six months of a death. Several online database ancestry and

genealogy websites incorporate the SSDI search engine, so researchers should take care to use only the native search site. We did not evaluate the engine at these secondary sites and cannot comment on their accuracy. Although our syncope study recorded accurate SSNs, this may not be the case for other ED-based studies, with a high proportion of undocumented subjects or frequent recording errors in demographic information.

CONCLUSION

Searching the online SSDI, with a correct SSN provides an accurate method to determine death as an outcome in clinical research studies. A majority of subjects ordinarily lost to follow-up can have their vital status determined.

Address for correspondence: James Quinn, MD, MS
Division of Emergency Medicine, Stanford University, 701 Welch Rd., Palo Alto, California 94304, Email: quinnj@stanford.edu

REFERENCES

1. Soteriades ES, Evans JC, Larson MG et al. Incidence and prognosis of syncope. *N Engl J Med* 2002; 347:878-885.
2. Su D, Fairbanks R, Shah M. Accuracy of the social security death index and the national death index as sources of mortality information for outcomes research. *Acad Emerg Med* 2006; 13:S147.
3. Social Security Death Index [Web Site]. Available at: <http://www.death-records.net/ssdi.htm>.
4. Schisterman EF, Whitcomb BW. Use of the Social Security Administration Death Master File for ascertainment of mortality status. *Popul Health Metr* 2004; 2:2.
5. Sesso HD, Paffenbarger RS, Lee IM. Comparison of National Death Index and World Wide Web death searches. *Am J Epidemiol* 2000; 152:107-111.
6. Quinn JV, Stiell IG, McDermott DA, Sellers KL, Kohn MA, Wells GA. Derivation of the San Francisco Syncope Rule to predict patients with short-term serious outcomes. *Ann Emerg Med* 2004; 43:224-232.