Annals of Epidemiology 56 (2021) 40–46

Contents lists available at ScienceDirect

Annals of Epidemiology

Original article

Confirming mortality in a longitudinal exposure cohort: optimizing National Death Index search result processing

Ingrid Giesinger, MSc [a], Jiehui Li, MBBS, MS [a, *], Erin Takemoto, PhD, MPH [a], Robert M. Brackbill, PhD, MPH [a], James E. Cone, MD, MPH [a], Baozhen Qiao, PhD [b], Mark R. Farfel, ScD [a]

[a] Division of Epidemiology, New York City Department of Health and Mental Hygiene, World Trade Center Health Registry, Long Island City, NY
[b] New York State Department of Health, New York State Cancer Registry, Albany, NY

ARTICLE INFO

Article history:
Received 26 February 2020
Accepted 20 October 2020
Available online 23 October 2020

Keywords:
Mortality
Vital statistics
Algorithms
National Death Index
Process linkage output

ABSTRACT

Purpose: The National Death Index (NDI) is an important resource for mortality ascertainment. Methods selected to process NDI search results are rarely described in studies using linked data and can have an impact on resources and mortality ascertainment. We evaluate methods to process NDI search results among a 9/11-exposed cohort—the World Trade Center Health Registry (Registry).
Methods: We describe three approaches to process search results (NDI-recommended cutoff points [NDIc]; National Program of Cancer Registries [NPCR] algorithm, and modified National Institute of Occupational Safety and Health algorithm [mNIOSH]). We calculate percent agreement, positive predictive value, sensitivity, specificity, and quantify the burden of manual review to compare the approaches.
Results: Of 51,158 Registry enrollees submitted for linkage, 9449 enrollee-level and 17,909 record-level matches were identified. NPCR and mNIOSH were highly concordant (97.1%); more record pairs required manual review for mNIOSH (mNIOSH: 2.7% and NPCR: 1.8%). NDIc sensitivity was 82.9%, with differences observed by race and ethnicity (Asian: 74.4% and White: 86.1%).
Conclusions: NPCR algorithm minimized false matches and reduced the manual review burden. NDIc had nonrandom distribution of missed matches and low sensitivity. NDI search processing methods have important implications for resulting linked data; measures of linkage quality should be available to data users.

This publication was supported by Cooperative Agreement Numbers 2U50/OH009739 and 5U50/OH009739 from the National Institute for Occupational Safety and Health (NIOSH) of the Centers for Disease Control and Prevention (CDC); U50/ATU272750 from the Agency for Toxic Substances and Disease Registry (ATSDR), CDC, which included support from the National Center for Environmental Health, CDC; and by the New York City Department of Health and Mental Hygiene. B.Q. is supported by the National Program of Cancer Registries of the Centers for Disease Control and Prevention through cooperative agreement 6NU58DP006309 awarded to the New York State Department of Health.

Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NIOSH, CDC, or the Department of Health and Human Services.

Ethics approval and consent to participate: The U.S. Centers for Disease Control and Prevention and the New York City Department of Health and Mental Hygiene Institutional Review Boards approved the Registry's protocol.

Consent for publication: not applicable.

Availability of data and material: The Registry owns the data and may make it available following review of applications to the Registry from external researchers.

Competing interests: The authors declare no competing interests.

Authors' contributions: I.G. and J.L. conceived and designed the study. I.G. performed the statistical analysis and drafted the article with input from coauthors. All authors contributed to data interpretation, revision of the article, and approved the final article.
* Corresponding author. New York City Department of Health and Mental Hygiene, World Trade Center Health Registry, 42-09 28th Street, CN-6W, Long Island City, NY 11101. Tel.: +1-718-786-4412.
E-mail address: jli3@health.nyc.gov (J. Li).

https://doi.org/10.1016/j.annepidem.2020.10.010
1047-2797/

Introduction

Comprehensive vital status ascertainment is an important component of many cohort studies. The National Death Index (NDI) compiles death certificate details from state and territorial vital statistics offices across the United States. NDI linkage is commonly used for mortality follow-up and cause of death determination in population-based disease surveillance or research studies. Following a researcher's submission of records with prescribed data elements [1], the NDI performs a linkage to the compiled death certificates and returns the search results; results can include more than one potential match per submitted record. To help researchers identify true matches, three indicators of the quality of match (probabilistic scores, class of the match, and match flag) are included.

Relying on the NDI indicators alone can result in the loss of true matches, especially when the initial NDI submission included incomplete data. To improve the accuracy in the determination of true matches, different algorithms have been developed to process NDI search results, including the Center for Disease Control and Prevention's (CDC) National Program of Cancer Registries (NPCR) algorithm [2]. The algorithm selected to process NDI results could have a potential impact on the assignment of vital status for certain search records and consequently affect research findings. For instance, missed matches reduce the number of deaths in the analytic sample, which artificially inflates the person time under observation [3,4]; differential misclassification of vital status by demographic factors (e.g., race, age) could alter the apparent survival rate of certain populations within the study [3,5]. Evaluating the quality of data linkage (sensitivity, specificity, etc.) can inform whether bias exists in the resulting linked data and subsequent analytical findings [4,6]. However, to our knowledge, no study has compared methods to process NDI search results considering measures of quality and bias in the resulting analytic study sample.

Evaluation of processing algorithms can improve processing efficiency, mortality surveillance, and data linkage in general. Transparency in methodology also assists the interpretation of findings from linked datasets, allows for reproducibility, identifies limitations, and informs comparability of findings between studies [7,8]. The World Trade Center Health Registry (Registry), a closed cohort of individuals exposed to the World Trade Center (WTC) disaster, adapted an algorithm developed by the National Institute for Occupational Safety and Health (mNIOSH) to process NDI search results for ongoing mortality surveillance [9,10]. To evaluate our current approach of processing search results, we compare the mNIOSH algorithm with the NPCR algorithm and the NDI-recommended cutoff allocation approach, examine the impacts on the resulting analytic sample, and share insight for other studies that use NDI linkage.

Material and methods

We first briefly describe our study population and mortality data source. We then describe three approaches to process the NDI search results: (1) NDI-recommended cutoff points (NDIc), (2) modified NIOSH algorithm (mNIOSH), and (3) NPCR algorithm (Fig. 1). We also describe a manual review process important to mNIOSH and NPCR and the quality measures applied to evaluate the approaches.

Study population

The Registry is a longitudinal cohort study of over 71,000 individuals enrolled after the WTC disaster [11,12]. Enrollment was conducted in 2003–2004, where personal identifiable information such as full name, date of birth, social security number (SSN), baseline information on demographics and health information, and exposures to the WTC disaster were collected.

The present study was limited to 51,158 enrollees whose identifying information was submitted to the NDI for linkage in 2018 to identify deaths that occurred during the period 2015–2016. Individuals who withdrew from the Registry, had known vital status, or had known insufficient identifying information were not included in the submission. The New York City Department of Health and Mental Hygiene Institutional Review Board approved the Registry's mortality protocol.

NDI linkage

The vital status of the Registry cohort has been updated periodically through linkage with NDI since 2008, a centralized nationwide database of mortality records at the U.S. National Center for Health Statistics, with the support of state vital statistics offices. After submission of an enrollee file to NDI, search results are returned to the user, often with enrollees matched to one or more NDI record. We use the term *record-level* to refer to the NDI death records involved in possible matches and the term *enrollee-level (enrollee)* to refer to the unique individual Registry enrollee whose record was submitted.

Three approaches used to process NDI linkage results

NDIc allocation

To determine a true match, NDIc relies on three indicators: (1) probabilistic score (the weighted sum of the probability of finding agreement between the NDI and user record), (2) class of the match (1–5, five mutually exclusive categories based on the number and the type of agreed identifying items where 1 is exact match), and (3) status code to flag true match (a binary allocation to "true" and "false" based on the probabilistic scoring) [1]. NDIc does not identify questionable matches for manual review [1]. Matches are considered "true" if they meet one of the following criteria: (1) matches with a class code = 1, (2) class code = 2 and probabilistic score >44.5, (3) NDI class code = 3 and probabilistic score >37.5, or (4) NDI class code = 4 and probabilistic score >32.5 [1].

mNIOSH algorithm

In 2010, we adapted and modified a NIOSH-provided SAS algorithm to process the NDI search results. The algorithm evaluates the search results using conditional logic based on the level of agreement between Registry enrollee data and NDI death record, all three NDI probabilistic scoring elements, and name probability. The probability of finding a particular combination of first and last names is provided by NDI and based on the frequency of the name in NDI records (1979–2005). The overall probability of a name was used as a cutoff point. This assumed independence of first and last names and required an exact first and last name match. In addition to the classification of "true" (appropriate matches which are retained) and "false" (inappropriate matches which are discarded), the algorithm includes allocation to "questionable," a gray zone [13,14], which requires further allocation to "true or false" through manual review. To reduce the likelihood of false and missed matches, we modified the NIOSH algorithm, increasing the quantity requiring manual review.

National Program of Cancer Registries algorithm (NPCR, reference algorithm)

NPCR, which has been developed and improved over time, is used by state cancer registries for cancer surveillance and has demonstrated high sensitivity and specificity [2]. NPCR classifies

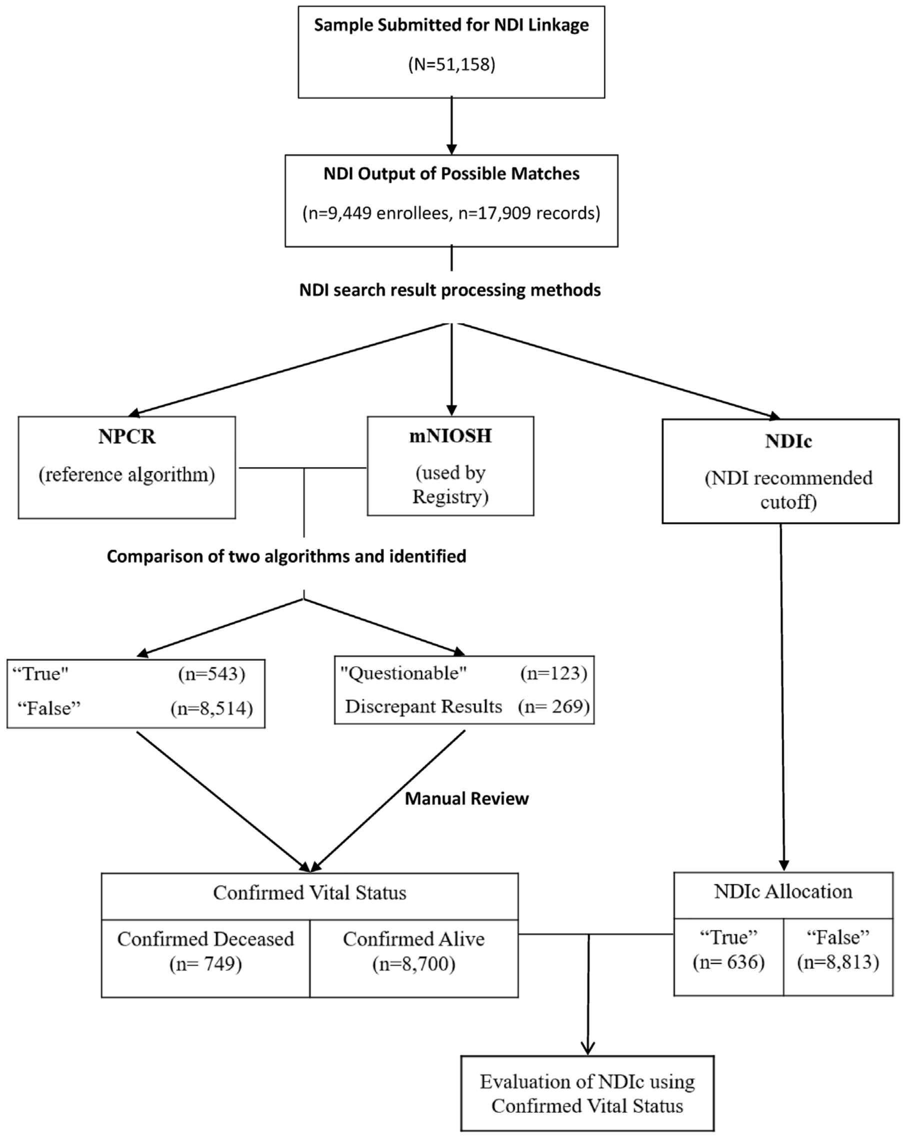

Fig. 1. Flowchart of mortality ascertainment and evaluation.

the NDI search results into three categories ("true," "false," or "questionable"). Using conditional logic, the algorithm assigns a "pass value" of 1–16 for the matches with specific agreements of data elements between enrollee and NDI records. Unlike mNIOSH, the NDI probabilistic score is not considered in the assignment of "pass value." Name frequency, a six-level categorical variable reflecting the rarity of the name based on NDI provided frequency, is included in the assignment of pass values along with consideration of other enrollee data elements. Those with a pass value of $\leq$4.3 and those with a pass value of 4.4–13.0 and an NDI status of 1 are allocated to "true" without further review. All other matches which received a pass value are considered "questionable" and are manually reviewed. Matches that did not receive a pass value are assumed to be "false" matches.

Comparison and evaluation of NDI processing algorithms

Comparison of mNIOSH and NPCR algorithms

The mNIOSH and NPCR algorithms were compared in terms of match allocation to "true," "false," and "questionable." Concordance was defined as agreement in the algorithm allocation. Percent agreement was defined as the proportion of confirmed algorithm allocations among the total number of matches manually reviewed. Discrepant algorithm allocations were used to calculate percent agreement for mNIOSH and NPCR allocations. For each algorithm allocation ("true"/"false"), the number of confirmed matches was divided by the number allocated for manual review.

The proportion of record pairs or matches (pp) requiring manual review to confirm vital status was calculated by dividing the sum of record-level matches allocated as "questionable" by the total number of record-level matches in the sample; a lower pp value being a measure of efficiency [13,14]. The pp is considered alongside quality metrics; false and missed matches provide an assessment of uncertainty in the resulting analytic samples [13–15].

For quality assurance, a convenience sample of record pairs allocated as "false" by both the mNIOSH and NPCR algorithms ($n = 21$) and record pairs allocated as "true" by both algorithms ($n = 9$) were included in the manual review. Algorithm allocation was not available to manual reviewers. This process allowed us to verify each algorithm's performance in its assignment to "true or false" and provided assurance of the NPCR algorithm process.

Table 1
Characteristics of WTC Registry enrollees across stages of vital status ascertainment through NDI linkage

Characteristic	Submitted for NDI linkage		Involved in possible NDI matches		Questionable and discrepant manually reviewed matches*	
	n	%	n	%	n	%
Total number of enrollees	51,158		9449		300*	
SSN						
Complete SSN	30,939	60.5	5573	59.0	53	17.7
Missing SSN†	20,219	39.5	3876	41.0	247	82.3
Race and ethnicity‡						
Non-Latino White	30,232	59.1	4665	49.4	145	48.3
Non-Latino Black	6535	12.8	1545	16.4	52	17.3
Latino	7672	15.0	1873	19.8	51	17.0
Asian	4372	8.6	915	9.7	31	10.3
Multiracial	1065	2.1	192	2.0	7	2.3
Other	1282	2.5	259	2.7	14	4.7
Gender						
Male	30,148	58.9	5774	61.1	183	61.0
Female	21,009	41.1	3675	39.0	117	39.0
Non-Latino White, *N*	30,232		4665		145	
Complete SSN	19,167	63.4	2928	62.8	19	13.1
Missing SSN†	11,065	36.6	1737	37.2	126	86.9
All other racial and ethnic groups, *N*	20,926		4784		155	
Complete SSN	11,772	56.3	2645	55.3	34	21.9
Missing SSN†	9154	43.7	2139	44.7	121	78.1

* Subsample excluded from manual review ($n = 92$) with NDI probabilistic score <27 and NDI class 3 or 4.
† Includes a small proportion (0.34% at submission), which were submitted as last four digits of SSN with leading zeros.
‡ Asian includes Native Hawaiian and Pacific Islander, and other includes American Indian, Alaskan Native and unknown.

Manual review process

Matches allocated to “questionable” by either mNIOSH or NPCR were evaluated independently by two reviewers and reclassified as “true” or “false”; discrepancies were evaluated by a third reviewer. The reviewers used the internal Registry database, which maintains a historical log of contact between the Registry and enrollee, and in some cases, notification of enrollee death from next of kin. Where insufficient information was available, external resources including obituary and memorial internet searches were used to confirm or reject matches, as well as LexisNexis facilitated searches of public records. Matches identified as “true” by both the NPCR and mNIOSH algorithms and those allocated “true” following the manual review were *confirmed deceased*. Matches identified as “false” by both algorithms and those allocated “false” following manual review were *confirmed alive*.

Evaluation of NDIc

Although gold standard data with a known true match status is preferred, a subsample of reference data, which has been manually reviewed, can be used to test linkage algorithms [8]. Using the confirmed vital status as reference, we evaluated the allocation of the NDIc to “true or false.” Statistical measures of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated on the total sample and by enrollee data elements involved in NDI linkage [8]. All statistical analyses were performed using SAS software, Version 9.4 (SAS Institute, Cary, NC).

Results

Of 51,158 enrollees submitted for the NDI linkage, 9449 enrollees were possibly matched. Approximately 40% (3644) of enrollees were matched with more than one NDI record, resulting in 17,909 record-level matches. Nearly 60% of enrollees submitted for NDI linkage were White, 15% were classified as Latino, and 8.6% as Asian (Table 1). Most manually reviewed matches had missing SSN (82.3%). Of the manually reviewed matches with complete SSN, a higher proportion were non-White (21.9%).

Comparison of mNIOSH evaluation and NPCR algorithms

Concordance was 97.1% between the NPCR and mNIOSH algorithms’ enrollee-level allocations to “true” (5.7%, $n = 543$), “false” (90.1%, $n = 8514$), and “questionable” (1.3%, $n = 123$; Table 2). Of the 2.8% ($n = 269$) discrepant enrollee-level results, 182 were identified

Table 2
Distribution of two classification algorithms to process NDI linkage result and determine mortality status of Registry enrollees

Algorithm allocation	Record level ($n = 17{,}909$)		Enrollee level ($n = 9449$)	
	n	%	*n*	%
Both algorithms identified as “true”	872	4.9	543	5.7
Both algorithms identified as “false”	16,427	91.7	8514	90.1
Both algorithms identified as “questionable”	198	1.1	123	1.3
Discrepant results	412	2.3	269	2.8
NPCR-“questionable,” mNIOSH-“true”	118	—	77	—
NPCR-“questionable,” mNIOSH-“false”	10	—	10	—
NPCR-“true,” mNIOSH-“questionable”	67	—	36	—
NPCR-“false,” mNIOSH-“questionable” (manually reviewed)	54	—	54	—
NPCR-“False,” mNIOSH-“questionable” (not manually reviewed)*	163	—	92	—

* Subsample excluded from manual review because of NDI probabilistic score <27 and NDI Class 3 or 4.

Table 3
Agreement of manual review outcome with algorithm allocation for a sample of NDI linkage pairs

Algorithm allocation	Total sample, *N*	Total manually reviewed ($n = 330$)	Manual review outcome		Percent agreement* (n confirmed/n reviewed)
			Match	Nonmatch	
Both algorithms identified as "true"	543	9†	9	0	100 (9/9)
Both algorithms identified as "false"	8514	21†	0	21	100 (21/21)
Both algorithms identified as "questionable"	123	123	96	27	—
Discrepant results					
NPCR-"questionable," mNIOSH-"true	77	77	72	5	93.5 (72/77)
NPCR-"questionable," mNIOSH-"false"	10	10	1	9	90.0 (9/10)
NPCR-"true," mNIOSH-"questionable"	36	36	36	0	100 (36/36)
NPCR-"false," mNIOSH-"questionable"	54	54	1	53	98.1 (53/54)
NPCR-"false," mNIOSH-"questionable"	163‡	—	—	—	—

* % agreement: number confirmed as match or nonmatchdivided by number selected for manual review.
† Subsample of respective allocation for quality assurance.
‡ Excluded from manual review because of NDI probabilistic score <27 and NDI Class 3 or 4.

as "questionable" by mNIOSH, and 87 "questionable" were identified by NPCR. There were no instances where one algorithm allocated as "true" and the other as "false."

Agreement of manual review of "questionable" or discrepant with algorithm allocation

A total of 330 enrollee-level matches were manually reviewed, which included 30 concordant algorithm allocations ("true" $n = 9$ of 543; "false" $n = 21$ of 8514). Algorithm allocation was confirmed in this subset (Table 3). Among the 123 "questionable" identified by both algorithms, 78.0% ($n = 96$) were considered true matches after review. Among discrepant results, concordance was 93.5% in those with *NPCR-"questionable," mNIOSH-"true,"* and 100% in those with *NPCR-"true", mNIOSH-"questionable."*

NPCR allocated a lower proportion of record-level matches to manual review, compared with mNIOSH (NPCR *pp*: 1.8%; mNIOSH *pp*: 2.7%). In total, 749 enrollees were considered confirmed deceased, 543 evaluated as "true" by both algorithms, and 206 identified through manual review.

NDIc evaluation

The NDIc identified 636 enrollees as "true," of which 621 were concordant with the confirmed deceased. This approach had high specificity (99.8%) and PPV (97.6%; Table 4). However, it missed 17% of those confirmed deceased ($n = 128$ of 749 confirmed deceased), resulting in a sensitivity of 82.9%. Although the specificity was similar across demographic characteristics, the sensitivity and PPV were lower for those missing SSN compared with those with complete SSN and for each non-White racial group. There was little difference in NPV (NPV >98%) across demographic characteristics (data not shown).

Discussion

We compared our existing program mNIOSH with the NPCR algorithm for processing NDI search results, and we also evaluated the NDIc approach using the confirmed vital status. The NPCR algorithm was an improvement on our existing approach, minimizing false matches and reducing questionable matches requiring review. Given the volume and nonrandom distribution of missed matches, we found the NDIc approach, an inferior option for processing NDI matches.

The NPCR algorithm effectively processes a high volume of NDI search results for state cancer registries and has been successfully modified to evaluate NDI linkage in a cohort of U.S. service members [2], although it was not modified in this evaluation. The algorithm reduced the number of records requiring manual review by 156 and returned fewer false matches compared with mNIOSH, demonstrating its effectiveness in at least three disparate study populations [2]. It is possible that customized revisions to the NPCR

Table 4
Estimates of linkage quality for NDIc classification compared with confirmed deceased*

Characteristic	Total	Confirmed deceased* ($n = 749$)		Confirmed alive ($n = 8700$)		Sensitivity, % (95% CI)	Specificity† (%)	PPV, %, (95% CI)
		NDIc	Not in NDIc	NDIc	Not in NDIc			
Total	9449	621	128	15	8685	82.9 (80.2, 85.6)	99.8	97.6 (96.1, 98.7)
SSN								
Complete SSN	5573	408	72	1	5092	85.0 (81.5, 88.1)	100.0	99.8 (98.7, 100.0)
Missing SSN	3876	213	56	14	3593	79.2 (73.8, 83.9)	99.6	93.8 (89.8, 96.6)
Race and ethnicity‡								
Non-Latino White	4665	404	65	2	4194	86.1 (82.3, 89.1)	100.0	99.5 (98.2, 99.9)
All other racial and ethnic groups	4784	217	63	13	4491	77.5 (72.1, 82.2)	99.7	94.4 (90.5, 97.0)
Non-Latino Black	1545	88	25	2	1430	77.9 (69.1, 85.1)	99.9	97.8 (92.2, 99.7)
Latino	1873	68	19	2	1784	78.2 (68.0, 86.3)	99.9	97.1 (90.1, 99.7)
Asian	915	29	10	7	869	74.4 (57.9, 87.0)	99.2	80.6 (64.0, 91.2)
Multiracial	192	17	5	0	170	77.3 (57.6, 92.2)	100.0	100.0 (80.5, 100.0)
Other	259	15	4	2	238	79.0 (54.4, 94.0)	99.2	88.2 (64.6, 98.5)
Gender								
Male	5774	394	80	11	5289	83.1 (79.4, 86.4)	99.8	97.3 (95.2, 98.6)
Female	3675	227	48	4	3396	82.6 (77.5, 86.8)	99.9	98.3 (95.6, 99.5)

* Among 9449 enrollees as possible matches provided by NDI, 749 deaths were confirmed deceased by NPCR, mNIOSH and manual review and considered as reference data when compared with matches from NDIc.
† 95% confidence intervals (CIs) are not shown as specificity is comparable across groups.
‡ Asian includes Native Hawaiian and Pacific Islander, and other includes American Indian, Alaskan Native and unknown.

algorithm could further optimize performance for other studies. The NPCR algorithm relies heavily on SSN; studies with more complete data on SSN would likely have a minimal burden of manual review with this approach.

The NDIc is an attractive approach, as it completely removes manual review. In the evaluation, the NDIc had a high specificity and relatively low false match rate; however, it missed a substantial number of those confirmed deceased, with demonstrated lower sensitivity among those with missing SSN and non-White enrollees. Missed matches lead to an overestimate of person time in cohort studies and an underestimation of deaths [3,4]. NDIc may be an appropriate choice where missingness on matching variables is minimal or in large studies where manual review is not feasible; however, potential bias should be considered in subsequent analyses.

Our findings highlight important lessons for those using linked data for research purposes by demonstrating differences in the confirmed deaths dependent on the processing approach. Consistent methodology to process NDI search results within and across sites in multicenter studies is critical to the validity and generalizability of mortality studies. The scientific community has promoted increased transparency for the purposes of research reproducibility, including data processing and linkage [7]. In addition to reproducibility, transparency also improves interpretation, especially among disparate findings from studies using the same or similar populations. Differences errantly attributed to reasons of theoretical or biological importance may sometimes simply arise because of different processing approaches.

Next, although we recommend the NPCR algorithm, its heavy reliance on SSN to generate and evaluate matches [16] is another important consideration for researchers using the resulting linked data. In our study, missing SSN was highest among those of minority race compared with non-Latino White enrollees. This differential missingness introduces the possibility of differential linkage rates and error among race-ethnic subgroups [1,8,17]. Although it is not possible to know from our study if these differences will introduce bias into the findings of the future research or if they simply reflect a bias in the linkage process, they are a limitation of using linked data of which the implications must be noted.

Certain steps can be taken to help compensate for SSN missingness among race–ethnic subgroups. First, the NDI submission allows for a more specific classification of individuals identifying as Asian (e.g., Chinese, Japanese, and Korean) [18]. Additional efforts to classify individuals into more specific groups using self-reported demographic data (preferred language, country of birth, etc.) may increase the likelihood of a match being identified by NDI. Second, names are used both in the initial NDI submission and in the processing algorithms. Asian and Latino naming conventions do not follow "Western" convention; using aliases reflecting cultural naming patterns may increase the likelihood of a match [19,20]. Finally, calculating the probability of name matching using frequency of occurrence in NDI linkage from 1979 to 2005 may not reflect the probability of a name in a more diverse sample like ours. Considering cultural diversity and naming practice, appropriate to the population under study, may further optimize data linkage, including NDI search result processing [21].

Limitations are noted. There was no "gold standard" dataset to evaluate the performance of the algorithms. Although the data set used to evaluate the NDIc was a nonrandom subsample of the NDI search results, it was comprised primarily of matches in the "gray zone." This subset is most prone to misclassification and therefore provides insight into the overall performance of the NDIc classification. Manual review of matches is subjective and will vary by institution dependent on the resources available. The manual review in this study was double blind, performed by two reviewers with consistent resources and discrepancies were reviewed by a third individual, minimizing errors in allocation.

Conclusions

The use of the NPCR algorithm improves efficiency compared with the current mNIOSH approach in use at the Registry, with a reduction in the number of uncertain matches that required manual review and a small reduction in false matches. The cost of manual review is justified to reduce the missed match rate, especially given the nonrandomness of classification errors. Data linkage is a powerful research tool, but measures of linkage quality should be available to those conducting analyses with the linked data as transparency in methodology improves interpretation and reproducibility of findings.

Acknowledgments

The authors thank the World Trade Center Health Registry enrollees for their participation; those involved in establishing the initial Registry methodology to process NDI search results, specifically Hannah Jordan; Janette Yung for her role in assisting with our methodology and manual review process; the Registry Panel Maintenance and IT support, specifically Lennon Turner and Jorge Guengue, for their assistance to facilitate the manual review process. The authors also thank Amy Kahn from the NYS Department of Health for her critical review of the article and Maria Schymura from the NYS Department of Health for her support of this project and the National Program of Cancer Registries for sharing their methodology to process NDI search results.

References

[1] NCHS, N.C.f.H.S.. National death index. https://www.cdc.gov/nchs/ndi/index.htm. [Accessed 14 December 2018].
[2] Skopp NA, Smolenski DJ, Schwesinger DA, Johnson CJ, Metzger-Abamukong MJ, et al. Evaluation of a methodology to validate National Death Index retrieval results among a cohort of U.S. service members. Ann Epidemiol 2017;27(6):397–400.
[3] Harron K. Introduction to data linkage. In: London: Administrative Data Research Network; 2016.
[4] Schmidlin K, Clough-Gorr KM, Spoerri A, Egger M, Zwahlen M. Impact of unlinked deaths and coding changes on mortality trends in the Swiss National Cohort. BMC Med Inform Decis Mak 2013;13:1.
[5] Bohensky MA, Jolley D, Sundararajan V, Evans S, Pilcher DV, Scott I, et al. Data linkage: a powerful research tool with potential problems. BMC Health Serv Res 2010;10:346.
[6] Greenland S, Lash TL, Rothman K. Modern epidemiology, Vol. 2008. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.
[7] Gilbert R, Lafferty R, Hagger-Johnson G, Harron K, Zhang L-C, Smith P, et al. GUILD: GUidance for Information about linking data sets. J Public Health (Oxf) 2018;40(1):191–8.
[8] Harron KL, Doidge JC, Knight HE, Gilbert RE, Goldstein H, Cromwell DA, et al. A guide to evaluating linkage quality for the analysis of linked data. Int J Epidemiol 2017;46(5):1699–710.
[9] Jordan HT, Brackbill RM, Cone JE, Debchoudhury I, Farfel MR, Greene CM, et al. Mortality among survivors of the Sept 11, 2001, World Trade Center disaster: results from the World Trade Center Health Registry cohort. Lancet 2011;378(9794):879–87.
[10] Jordan HT, Stein CR, Li J, Cone JE, Stayner L, Hadler JL, et al. Mortality among rescue and recovery workers and community members exposed to the September 11, 2001 World Trade Center terrorist attacks, 2003-2014. Environ Res 2018;163:270–9.
[11] Brackbill RM, Thorpe LE, DiGrande L, Perrin M, Sapp JH, Wu D, et al. Surveillance for World Trade Center disaster health effects among survivors of collapsed and damaged buildings. MMWR Surveill Summ 2006;55(2):1–18.
[12] Farfel M, DiGrande L, Brackbill R, Prann A, Cone J, Friedman S, et al. An overview of 9/11 experiences and respiratory and mental health conditions among World Trade Center Health Registry enrollees. J Urban Health 2008;85(6):880–909.
[13] Christen P, Goiser K. Quality and complexity measures for data linkage and deduplication. In: Guillet FJ, Hamilton HJ, editors. Quality measures in data mining. Berlin, Heidelberg: Springer Berlin Heidelberg; 2007. p. 127–51.
[14] Gu L, Baxter R. Decision models for record linkage. In: Williams GJ, Simoff SJ, editors. Data Mining. Berlin, Heidelberg: Springer; 2006. p. 146–60.

[15] Harron K, Dibben C, Boyd J, Hjern A, Azimaee M, Barreto ML, et al. Challenges in administrative data linkage for research. Big Data Soc 2017;4(2). 2053951717745678.
[16] Sayer B, Cox C. How many digits in a handshake? National Death Index matching with less than nine digits of the Social Security Number. In: Joint Statistical Meetings - Section on Survey Research Methods. San Francisco: American Statistical Association; 2003.
[17] Cowper DC, Kubal JD, Maynard C, Hynes DM. A primer and comparative review of major US mortality databases. Ann Epidemiol 2002;12(7):462–8.
[18] Thompson CA, Gomez SL, Hastings KG, Kapphahn K, Yu P, Shariff-Marco S, et al. The burden of cancer in Asian Americans: a report of National Mortality Trends by Asian ethnicity. Cancer Epidemiol Biomarkers Prev 2016;25(10):1371–82.
[19] Tibble H, Law HD, Spittal MJ, Karmel R, Borschmann R, Hail-Jares K, et al. The importance of including aliases in data linkage with vulnerable populations. BMC Med Res Methodol 2018;18(1):76.
[20] Larney S, Burns L. Evaluating health outcomes of criminal justice populations using record linkage: the importance of aliases. Eval Rev 2011;35(2):118–28.
[21] Shlomo N. Overview of data linkage methods for policy design and evaluation. In: Data-driven policy impact evaluation. Cham, Switzerland: Springer; 2019. p. 47–65.