American Journal of Epidemiology

Vol. 174, No. 6
DOI: 10.1093/aje/kwr155
Advance Access publication:
July 29, 2011

Practice of Epidemiology

Imputation of Incident Events in Longitudinal Cohort Studies

George Howard*, Leslie A. McClure, Claudia S. Moy, Monika M. Safford, Mary Cushman, Suzanne E. Judd, Brett M. Kissela, Dawn O. Kleindorfer, Virginia J. Howard, David J. Rhodes, Paul Muntner, and Hemant K. Tiwari

* Correspondence to Dr. George Howard, Department of Biostatistics, School of Public Health, University of Alabama at Birmingham, Ryals Building, Room 327D, 1665 University Boulevard, Birmingham, AL 35294-0022 (e-mail: ghoward@uab.edu).

Initially submitted May 5, 2010; accepted for publication April 15, 2011.

Longitudinal cohort studies normally identify and adjudicate incident events detected during follow-up by retrieving medical records. There are several reasons why the adjudication process may not be successfully completed for a suspected event including the inability to retrieve medical records from hospitals and an insufficient time between the suspected event and data analysis. These "incomplete adjudications" are normally assumed not to be events, an approach which may be associated with loss of precision and introduction of bias. In this article, the authors evaluate the use of multiple imputation methods designed to include incomplete adjudications in analysis. Using data from the REasons for Geographic And Racial Differences in Stroke (REGARDS) Study, 2008–2009, they demonstrate that this approach may increase precision and reduce bias in estimates of the relations between risk factors and incident events.

cohort studies; imputation; longitudinal studies; missing data

Abbreviation: REGARDS, REasons for Geographic And Racial Differences in Stroke.

It is common that ongoing longitudinal cohort studies publish associations of outcomes with exposures while follow-up continues. For example, the Framingham Study originally reported the association of blood pressure and stroke risk in 1970 (1) and has updated this description in reports including 1978 (2), 1981 (3), 1991 (4), 1994 (5), 2001 (6), and 2008 (7). Publishing results during acquisition of data introduces the need for ongoing surveillance and adjudication for events, with an associated ongoing process of data closure, analysis, and reporting of results.

The surveillance and adjudication process for incident events in longitudinal cohort studies is a complex process. Although study methods differ, suspected incident events usually result in a request for medical records for evaluation by an adjudication committee. Study personnel may have to contact the study participant to provide updated permission for medical record release, a request that can be honored or declined. If permission is given, the hospital is asked to release the associated information. On receipt of this information, packages are assembled for the adjudication committee.

Failure to reach definitive adjudications can occur for many reasons. Despite participants providing initial consent for medical record retrieval, they may not provide updated release. Even with a signed medical release form, the hospital may not be willing or able to provide information or to locate the information. Therefore, in all studies, there are records that cannot be retrieved to document the details of a suspected event.

Once received, the time to process records can be quite heterogeneous. When data sets are closed for analysis, an arbitrary time in the past is frequently chosen to permit time to process events. As most records can be quickly adjudicated, the selection of a recent time will increase the number of observed events in the analysis, thereby improving precision of estimates. However, selecting a recent time implies a larger number of events are still in process. Hence, choosing the time for data closure is not a straightforward decision.

There are 2 primary types of missing events: those where medical records cannot be retrieved and those where suspected events are still in the medical record retrieval/adjudication process. We refer to either of these as ''incomplete adjudication,'' while ''complete adjudications'' refer to those with definitive decisions by the adjudication committee (regardless of adjudication outcome). To our knowledge, previous studies have not considered the impact of incomplete adjudications, therefore implicitly assuming these are not events. However, if risk factors are associated with the ability to complete an adjudication, ignoring incomplete adjudications may introduce bias in the assessment of the association of the risk for these same risk factors. For example, if records from individuals with lower socioeconomic status are more difficult to retrieve, then the proportion of events with incomplete adjudications will be greater, resulting in the underestimation of risk in lower socioeconomic status individuals.

Incomplete adjudications can be considered as missing data. There have been recent calls to increase the use of multiple imputation techniques to improve precision and reduce bias when data are missing (8). We propose use of multiple imputation using 2 approaches:

- Approach 1. Select a time for data closure far enough in the past to ensure adequate time for adjudication of most events and impute outcomes only where records could not be retrieved. Here, the suspected events in process at the time of the data closure are not counted as cases. This approach has the advantage of requiring imputation only for failure to retrieve the record but the disadvantage of failing to include recently adjudicated events.
- Approach 2. Choose a more recent time and impute outcomes from both records with retrieval failure and those still in process. This approach has the advantage of including all completely adjudicated events but the disadvantage of needing to impute a larger number of events for those still in process.

The impact of imputation and the choice between these alternative approaches is considered by using data from the REasons for Geographic And Racial Difference in Stroke (REGARDS) Study.

MATERIALS AND METHODS

REGARDS is a longitudinal cohort study evaluating the causes of racial and geographic differences in stroke mortality in the United States. The study is following 30,239 African-American and white participants aged 45 years or older recruited from 1,834 US counties. Twenty-one percent (21%) of participants were from the ''buckle'' of the Stroke Belt along the coastal plain region of North Carolina, South Carolina, and Georgia; 35% were from the Stroke Belt states comprising the remainder of North Carolina, South Carolina, and Georgia, plus Alabama, Mississippi, Tennessee, Arkansas, and Louisiana; and 44% were from the other 40 contiguous US states. Individuals were selected from commercially available lists of residents and contacted by using mail and telephone. For those agreeing to participate, self-reported risk factor information was obtained by computer-assisted telephone interview. Physical measures were collected at an in-home examination including blood pressure, blood and urine samples, and electrocardiogram. Participants were followed at 6-month intervals by telephone for surveillance including stroke events. An investigation of a suspected stroke is generated if a participant reports a stroke, transient ischemic attack, or hospitalization for stroke-like symptoms. Medical records are sought for all suspected strokes and processed through an adjudication committee. Records are first reviewed by a neurologic nurse specialist to eliminate records that were clearly nonstroke, with all potentially stroke-related records advancing to physician review. Records are examined by 2 physician reviewers, and the process is terminated if there is agreement. If there is not agreement, the case is discussed by the entire adjudication committee until consensus is achieved. Study methods were reviewed and approved by the institutional review boards at the collaborating institutions. Additional methodological details are provided elsewhere (9).

We implemented a multiple imputation approach logistic regression to establish predictors of stroke events among complete adjudications, and then these estimated the likelihood of events for imputation for the incomplete adjudications. For approach 1, requiring selection of a date in the past for data closure, we divided the study follow-up into calendar quarters and set the data closure at the closest calendar quarter with more than 80% of adjudications completed. Records still in process for events after the selected closure date were ignored, imputing outcomes only for suspected strokes where records were not retrieved. For approach 2, we used adjudications completed through the analysis date and imputed events where we were unable to retrieve medical records and where records were still in the adjudication process.

For both approaches, the imputation approach requires the joint processing of suspected events and participants (Figure 1). First, the suspected events (3,998 for approach 1/4,516 for approach 2, hereafter shown as ''3,998/4,516'' for the 2 approaches) are divided into those with complete (3,522/3,681) and incomplete (476/835) adjudications. With the complete adjudications, a logistic model was developed predicting a positive event adjudication. Predictors considered for inclusion in the model were race, sex, and region (a priori included because of importance to study goals); the reason the record was sought (reported stroke, reported transient ischemic attack, reported stroke symptoms, death report where stroke was suspected, reported myocardial infarction with a suspected stroke, hospitalization with a suspected stroke, incapacitation for any reason, and miscellaneous reasons that may include stroke); age (considered in 10-year intervals); and a summary measure of stroke risk factors (age- and sex-adjusted residual of the Framingham Stroke Risk Function (5) that incorporates hypertension and hypertension treatment, diabetes, cigarette smoking, atrial fibrillation, left ventricular hypertrophy, and history of heart diseases). For the factors potentially considered as predictors (those other than race, sex, and region), backward stepwise regression was used to remove factors that were not associated ($P > 0.05$) with the likelihood of a positive adjudication. The coefficients from this model were then used to inform the prediction of the probability that each incomplete adjudication would be positively adjudicated as a stroke event, and then these probabilities were used to create 10

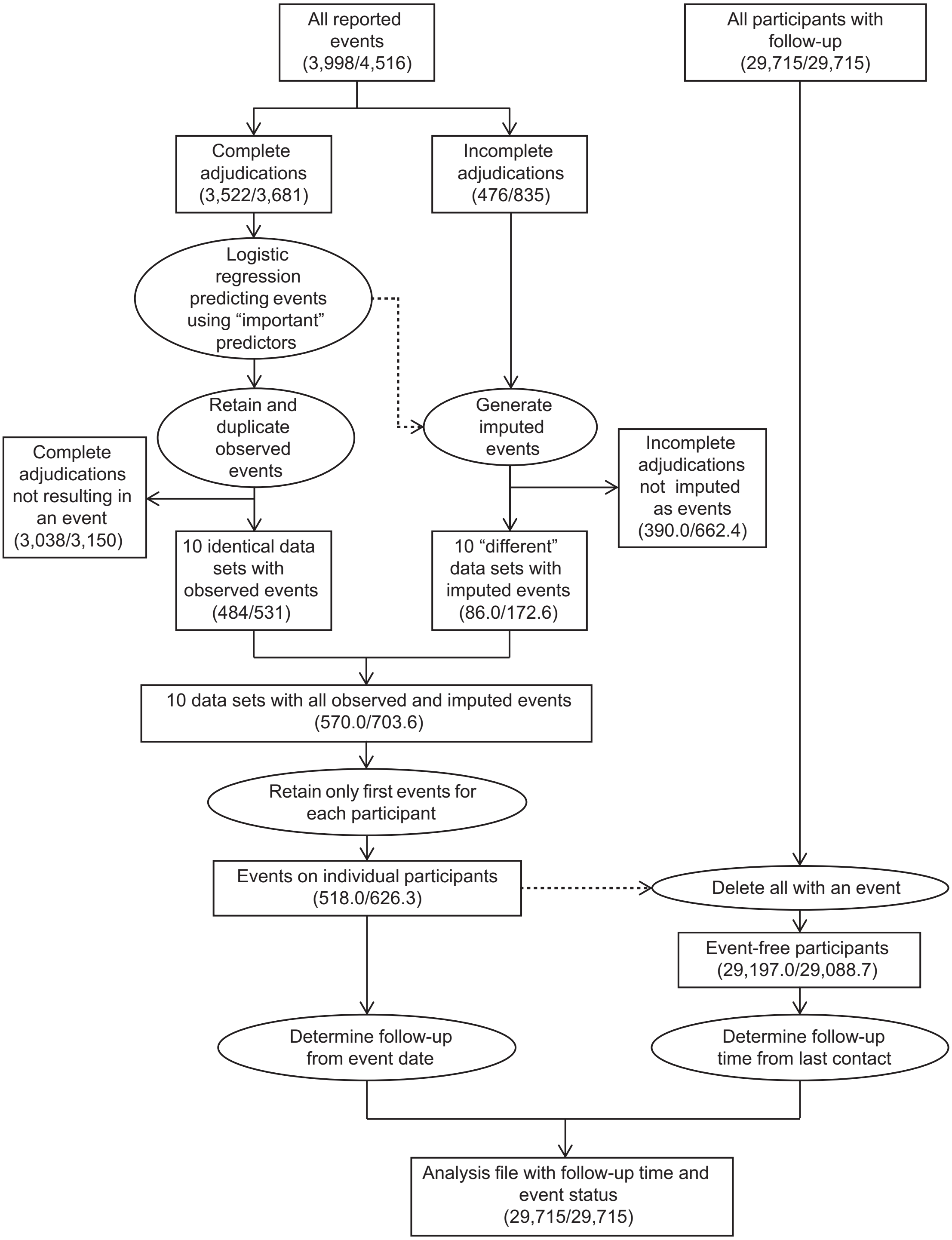

Figure 1. Flow diagram for processing suspected events and participants in the creation of the 10 imputation data sets and indicating the number of events or the number of participants at each step (approach 1/approach 2), REGARDS Study, 2008–2009. REGARDS, REasons for Geographic And Racial Differences in Stroke.

data sets with events imputed from the binomial distribution (on average, 86.0/172.6 events per data set). Those incomplete adjudications not imputed as events (an average of 390.0/662.4) did not continue in this process, while those that were imputed to be positive for an event did continue in the process.

Similarly for complete adjudications, those where the adjudication did not result in a positive event were discarded

(3,038/3,150), while those that were positively adjudicated continued in the process (484/531). For these, 10 identical data sets were created (each corresponding to 1 of the 10 imputed data sets).

Ten data sets with all positive events (an average of 570.0/703.6 events per data set) were then created by combining the 10 imputed event and 10 observed event data sets. These were then sorted by event date, and a single record was retained corresponding to the first event for the participant (518.0/626.3 participants per data set), now moving from processing of events to processing of participants.

The outcome for all participants without an event was then created by considering all participants for whom we had follow-up in the REGARDS Study (29,715/29,715). Ten data sets were created by removing all participants positive for an event in the process above, resulting in an average of 29,197.0/29,088.7 event-free participants in each data set. These 10 data sets contain no records for participants with a positively adjudicated event (484/531), and they contain identical records for the majority of participants without a suspected stroke event (29,715 − 3,998 = 25,717/29,715 − 4,516 = 25,199) and differ only by the inclusion or exclusion of participants with/without an imputed event. For each individual in each data set, an indicator was set to show the participant to be event free, and the follow-up time was set to the time of the last contact.

Records for those with/without events were combined to produce the 10 imputation data sets, each with 29,715 REGARDS participants with follow-up, with an indicator of event status and the follow-up time.

Poisson regression models were used to estimate the age- and sex-adjusted stroke event rates and incident rate ratios by race and region, with the estimate and its standard error being calculated from the 10 individual analyses by SAS PROC MIANALYZE (SAS Institute, Inc., Cary, North Carolina). Four separate analyses were performed: observed and observed + imputed outcomes were calculated for both approaches.

RESULTS

Although the closure date for approach 1 is arbitrary, we set it for quarters with at least 80% of reported events processed. This rule excluded data from the final quarter of 2008 (in which 78% of suspected events were adjudicated) and all of 2009 (with 58% adjudicated in quarter 1, 36% in quarter 2, and 23% in quarter 3). By use of this rule, 2.7% of all suspected strokes were in process, and 9.8% of all events adjudicated as stroke occurred after this closure date and were excluded. For approach 1, 3,998 suspected strokes in 3,017 individuals were identified. For approach 2, all data through the date of the last adjudication meeting (October 1, 2009) were used. In this approach, events that could not be retrieved and events in process were included in the imputation. Here, record retrieval had been attempted for 4,518 suspected strokes for 3,327 individuals. Records for two individuals who subsequently dropped out of the study were deleted, leaving 4,516 suspected strokes in 3,325 participants.

Table 1 provides a description of suspected events shown by the reason prompting retrieval and the status of the adjudication for the 3,998 suspected events for approach 1 and the 4,516 for approach 2. For approach 1, adjudication was complete for 3,522 (88.1%; 484 stroke events + 3,038 nonstrokes) of suspected strokes, and 476 (11.9%) records could not be retrieved and, thus, were incomplete adjudications. For approach 2, adjudication was complete for 3,681 of suspected strokes (81.5%; 531 stroke events + 3,150 nonstrokes); 305 (6.8%) suspected strokes were incomplete because they were still in process, and 530 (11.7%) were incomplete because records could not be retrieved.

Some records sought as suspected strokes did not contribute to the imputation process (Table 1). For myocardial infarction events with suspected strokes, none resulted in stroke outcome, and as such none of the incomplete adjudications would be imputed as a stroke. In addition, there were no incomplete adjudications pending for records sought for miscellaneous purposes and, as such, this category did not contribute to the imputations. Likewise there were no incomplete adjudications among deaths, and as such this group did not contribute to imputations for approach 1 (but 5 records still in process did contribute to imputations in approach 2).

As noted, if the likelihood of a completed adjudication differs by risk factors, then bias may be introduced in estimates of risk. For approach 2, considering all suspected strokes, African-American participants were more likely to have incomplete adjudications with 190 of 855 (22.2%) suspected strokes for African-American men and 295 of 1,385 (21.3%) for African-American women, compared with 181 of 1,255 (14.4%) for white men and 169 of 1,021 (16.6%) for white women. The likelihood of an incomplete adjudication did not differ by region ($P = 0.32$), with 287 of 1,327 (17.8%) from the Stroke Belt, 152 of 856 (17.8%) from the stroke buckle, and 396 of 2,046 (19.4%) from the rest of the nation incomplete. Similar patterns were observed among the suspected strokes considered for approach 1, but details were not provided for brevity.

Two logistic risk functions were developed relating the likelihood of a stroke outcome to stroke risk factors among the 3,522 completed adjudications in approach 1 and 3,681 completed adjudications in approach 2. Conditional on a report of a suspected triggering of a stroke adjudication, the stroke risk factors and age were not associated ($P > 0.3$) with the likelihood of being adjudicated as a stroke; however, there was a strong relation ($P < 0.0001$; c statistic 0.854 for both approaches) for the reason triggering record retrieval. Because of the importance of region, race, and sex to the current study, these factors were maintained in the logistic function despite not being significantly related to the likelihood of being adjudicated as stroke (all P's > 0.05). The odds ratios (and 95% confidence intervals) from the final model predicting confirmation of stroke among complete adjudications are provided in Table 2.

For each approach, 10 imputed data sets were generated. As shown in Table 3 (showing the number of *participants* with observed and imputed events), this resulted in an average of 518.0 total stroke events for approach 1 and 626.3 total stroke events for approach 2, of which 13.4% and 22.6%, respectively, were imputed. For both approaches, over half of the imputed events were for records retrieved for reported strokes (59% for approach 1 and 54% for approach 2) or

Table 1. Status of Records on Suspected Events That Were Pursued as Part of the Stroke Adjudication Process, REGARDS Study, 2008–2009[a]

	Approach 1: Data up to October 1, 2008							Approach 2: Data up to October 1, 2009								
	No. of Records	Complete Adjudications				Incomplete Adjudications (Record not Retrieved)		No. of Records	Complete Adjudications				Incomplete Adjudications			
		Stroke		No Stroke					Stroke		No Stroke		In Process		Record not Retrieved	
		No.	%	No.	%	No.	%		No.	%	No.	%	No.	%	No.	%
All	3,998	484	12.1	3,038	76.0	476	11.9	4,516	531	11.8	3,150	69.8	305	6.8	530	11.7
Reason for record retrieval																
Reported stroke	827	269	32.5	424	51.3	134	16.2	985	296	30.1	454	46.1	84	8.5	151	15.3
Reported transient ischemic attack	668	118	17.7	461	69.0	89	13.3	835	129	15.4	507	60.7	99	11.9	100	12.0
Stroke symptom	2,061	26	1.3	1,848	89.7	187	9.1	2,155	28	1.3	1,862	86.4	69	3.2	196	9.1
Deaths with a suspected stroke	39	25	64.1	14	35.9	0	0.0	50	30	60.0	15	30.0	5	10.0	0	0.0
Reported myocardial infarction with stroke suspicion	12	0	0.0	11	91.7	1	8.3	15	0	0.0	13	86.7	1	6.7	1	6.7
Hospitalized for other reasons	304	13	4.3	255	83.9	36	11.8	356	13	3.7	272	76.4	30	8.4	41	11.5
Incapacitation	76	26	34.2	21	27.6	29	38.2	108	27	25.0	23	21.3	17	15.7	41	38.0
Miscellaneous	11	7	63.6	4	36.4	0	0.0	12	8	66.7	4	33.3	0	0.0	0	0.0

Abbreviation: REGARDS, REasons for Geographic And Racial Differences in Stroke.

[a] Data are shown for "approach 1," where the data are closed at some (arbitrary) point in the past (for this table, October 1, 2008) and events are imputed only for incomplete adjudications with failure to retrieve information, and "approach 2," where very recent data are used (October 1, 2009) with data for incomplete adjudications for either failure to retrieve information or because the adjudication is still in process.

Table 2. Odds Ratios and 95% Confidence Intervals for Logistic Function Predicting the Likelihood of a Stroke Event From the Adjudication Process (Shown for Both Approaches), REGARDS Study, 2008–2009

Factor	Approach 1: Data up to October 1, 2008		Approach 2: Data up to October 1, 2009	
	OR	95% CI	OR	95% CI
Region (referent = rest of nation)				
Stroke belt	0.95	0.73, 1.22	0.98	0.77, 1.24
Stroke buckle	1.32	0.97, 1.78	1.24	0.93, 1.65
Race-gender (referent = white female)				
African-American female	1.05	0.76, 1.46	0.98	0.72, 1.33
African-American male	1.53	1.08, 2.15	1.37	0.99, 1.90
White male	1.26	0.93, 1.71	1.16	0.88, 1.55
Reason for seeking record (referent = reported stroke symptom)				
Reported stroke	45.44	29.92, 69.02	43.38	29.02, 64.86
Reported transient ischemic attack	18.24	11.76, 28.28	16.83	11.04, 25.67
Death with a suspected stroke	N/A		41.52	19.77, 95.80
Hospitalized for other reasons	3.61	1.83, 7.14	3.15	1.61, 6.18
Incapacitated	86.14	42.94, 172.73	76.14	38.87, 149.17

Abbreviations: CI, confidence interval; N/A, not available; OR, odds ratio; REGARDS, REasons for Geographic And Racial Differences in Stroke.

transient ischemic attack (17% for approach 1 and 23% for approach 2).

Table 4 provides the estimated stroke incidence rates and incidence rate ratios for race and region (after adjustment for sex and age) by multiple imputation techniques. The use of imputation increased the incidence rate proportionately with the number of imputed events (approximately 13% in approach 1 and 21% in approach 2) and had little impact on the estimated incidence ratio between regions. However, there was a substantial impact on the estimated incidence ratio for race (African American compared with white).

DISCUSSION

With the standard methods in a longitudinal study, a participant cannot be classified as having an event without completion of the adjudication process. This approach assumes that both types of incomplete adjudications, those with failure to retrieve records and those in process, are noncases, as Samuel Shem notes in his novel *House of God*: "If you don't take a temperature, you can't find a fever" (11, p. 147).

Using data from the REGARDS Study, we illustrate how incomplete adjudications can affect event rate estimates and

Table 3. Participants With Events Observed and Imputed (Across the 10 Observed and Imputed Data Sets), REGARDS Study, 2008–2009

	Approach 1: Data up to October 1, 2008					Approach 2: Data up to October 1, 2009				
	No. of Events	Event Type				No. of Events	Event Type			
		Observed		Imputed			Observed		Imputed	
		No.	%	No.	%		No.	%	No.	%
All	518.0	448.8	86.6	69.2	13.4	626.3	485.0	77.4	141.3	22.6
Reason for record retrieval										
Reported stroke	292.4	251.8	86.1	40.6	13.9	349.4	273.3	78.2	76.1	21.8
Reported transient ischemic attack	121.1	109.0	90.0	12.1	10.0	151.9	118.7	78.1	33.2	21.9
Stroke symptom	27.7	25.0	90.3	2.7	9.7	28.6	26.0	90.9	2.6	9.1
Death	23.0	23.0	100.0	0	0.0	26.9	25.0	92.9	1.9	7.1
Hospitalized for other reasons	15.3	13.0	85.0	2.3	15.0	16.4	13.0	79.3	3.4	20.7
Incapacitation	32.5	21.0	64.6	11.5	35.4	46.1	22.0	47.7	24.1	52.3
Miscellaneous	6.0	6.0	100.0	0	0.0	7.0	7	100.0	0	0.0

Abbreviation: REGARDS, REasons for Geographic And Racial Differences in Stroke.

Table 4. Estimated Incidence Rates and Incidence Rate Ratios From Observed Data and With the Use of Imputation, Shown for Race and Region, Using Both Approach 1 and Approach 2, REGARDS Study, 2008–2009

	Approach 1: Data up to October 1, 2008								Approach 2: Data up to October 1, 2009							
	Observed		Imputed		Observed		Imputed		Observed		Imputed		Observed		Imputed	
	IR (per 100,000)	95% CI	IR (per 100,000)	95% CI	IRR	95% CI	IRR	95% CI	IR (per 100,000)	95% CI	IR (per 100,000)	95% CI	IRR	95% CI	IRR	95% CI
Region																
Nonbelt	413.2	338.0, 505.2	489.2	405.0, 590.8	1.0	Referent	1.0	Referent	383.5	317.0, 463.9	506.6	423.1, 606.5	1.0	Referent	1.0	Referent
Belt	437.7	352.1, 544.1	499.6	405.4, 615.5	1.06	0.86, 1.31	1.02	0.84, 1.25	418.5	340.9, 513.7	542.4	448.3, 656.3	1.09	0.89, 1.34	1.07	0.88, 1.30
Buckle	526.0	409.1, 676.2	614.7	484.7, 779.5	1.27	0.99, 1.63	1.26	0.99, 1.59	482.4	380.2, 612.1	638.8	509.1, 801.5	1.26	0.99, 1.60	1.26	1.01, 1.58
Race																
White	394.1	324.7, 487.4	441.4	366.1, 532.3	1.0	Referent	1.0	Referent	376.3	313.3, 451.2	464.9	391.4, 552.0	1.0	Referent	1.0	Referent
African American	528.8	430.9, 648.9	640.1	529.3, 774.0	1.34	1.11, 1.62	1.45	1.21, 1.74	482.7	397.3, 586.5	674.4	563.2, 807.5	1.28	1.07, 1.54	1.45	1.23, 1.71

Abbreviations: CI, confidence interval; IR, incidence rate; IRR, incidence rate ratio; REGARDS, REasons for Geographic And Racial Differences in Stroke.

estimates of relative risk. The likelihood of an incomplete adjudication was higher for African Americans than for whites; thus, a larger proportion of strokes are likely undetected in African Americans than in whites, which would bias the African-American/white incidence rate ratio downward. The use of imputation identified a larger number of undocumented strokes among those not retrieved for African Americans, resulting in an increase in the African-American/white incidence rate ratio from 1.34 and 1.28 (in approach 1 or approach 2, respectively) without imputation to 1.45 and 1.45 with imputation. Imputation had little impact on regional differences because the incomplete adjudication rate did not differ by region.

Few publications report success rates of medical record retrieval. In the Women's Health Initiative (WHI), the success of retrieving event records was 85.9% in 2009 (Cora E. Lewis, University of Alabama at Birmingham, personal communication, 2010). Although medical record retrieval rates are not routinely reported, our experience in other studies is similar. Of note, Women's Health Initiative participants were seen in clinics as part of the study, and staff had numerous telephone contacts with participants, providing opportunities to establish rapport. By contrast, a record retrieval rate of 88.2% in the REGARDS Study was achieved after only 1 in-home visit at baseline, followed by semiannual telephone interviews. It seems likely that higher success rates will be in studies with stronger relationships with participants.

We suggest that the choice between approach 1 (closing data in the past and imputing adjudications only for failure to retrieve information) and approach 2 (selection of a recent time and imputation of all incomplete adjudications) should be driven by the proportion of recently reported events in the review process. If this proportion is small, little is gained by selecting approach 2. In the REGARDS Study, closing the data approximately 1 year in the past reduced the proportion of events still in the adjudication process from 6.8% to 2.7%, but it also excluded 9.8% of all observed events—a tradeoff that does not appear wise. Other studies may wisely use approach 1 if there is a more rapid adjudication process and/or a lower proportion of recent suspected events.

Two substantial gains may be made by imputation of events for incomplete adjudications. First, a gain in statistical precision may occur, especially if a larger number of events is available by approach 2. Consider the gain in precision as indexed by the width of the 95% confidence interval for the estimated incidence rate ratio for the stroke buckle versus non-Stroke Belt regions (Table 4). Analysis of the data without imputation (i.e., the column for "approach 1, observed") would have resulted in an estimated incidence rate ratio of 1.27 for the stroke buckle versus non-Stroke Belt regions with a wide confidence interval of 0.64 ($1.63 - 0.99 = 0.64$). By use of imputation, approach 1 does not affect the estimated incidence ratio of 1.26 but marginally narrows the confidence interval to 0.60 ($1.59 - 0.99 = 0.60$). However, use of approach 2 had little impact on the estimated incidence rate ratio of 1.26 but further improved the precision as the width of the confidence interval was reduced to 0.57 ($1.58 - 1.01 = 0.57$).

Second, there may be a reduction in the bias of the estimated rate ratios. This is important if risk factors are related to incomplete adjudications. Many factors could influence the completeness and timeliness of record retrieval so it seems likely that, without imputation, estimates will be biased. Specifically, in the REGARDS Study, incomplete adjudications for suspected stroke events were more likely among African-American than white participants. This implies that a larger proportion of records were de facto classified as nonevents for African Americans, resulting in a substantially lower incidence rate ratio without imputation. We suggest that differences in incidence rate ratios are due to the bias introduced from higher incomplete adjudication rates for African Americans.

Despite greater precision of incidence estimates from surveillance studies, incidence estimates are commonly reported from longitudinal cohort studies (12, 13). Because an incomplete adjudication is equivalent to a negative outcome from adjudication, incidence estimates from longitudinal cohort studies without imputation may be biased. The magnitude of this bias can be reduced by limiting the number of incomplete adjudications; however, there will always be some incomplete adjudications, and imputation of these suspected events will reduce the associated bias.

The introduction of imputation as an alternative analysis approach raises the issue of whether a study (such as the REGARDS Study) should have a protocol to dictate an analysis approach, potentially limiting the privilege of a first author (and analyst) to determine the approach used in his/her manuscript. For the REGARDS Study, because of the known potential bias introduced by racial differences in the ability to retrieve records, the steering committee has decided to "very strongly encourage" the use of imputation for reports focusing on racial differences (or in papers where race is considered as an effect modifier). For papers where race is not emphasized, we propose to leave the choice of the analysis approach more open to the first author and analyst but do feel it important to make investigators aware of this approach.

There are clearly limitations of the approaches described herein; most notably, use of the technique presents one set of prevalence and relative risk estimates from observed data and a second set from the imputed data—raising the question of which (if either) is correct. Key to this answer is the assumption that the likelihood of an incident event is similar for incomplete and complete adjudications. Under this assumption, multiple imputation will provide unbiased estimates of both prevalence and relative risk and will be the correct answer. Although this assumption may be reasonable for the REGARDS Study, it is difficult to confirm its validity among those adjudication attempts where records could not be retrieved. We feel that the reasons for failing to retrieve the records (including challenges with specific hospitals, individuals declining to provide permission to retrieve records, and so on) are not likely to be associated differentially with the likelihood of a potential case being adjudicated as a stroke. There is also the opportunity to improve the approach in the step where logistic regression is used to predict likelihood of an incomplete adjudication being a stroke event by introducing a sampling from the multivariate normal distribution of the estimated logistic regression parameters to create a different predictive logistic regression model for each imputation data set. Technically, this step is required to ensure that the imputation is "proper" (10). In the case of the REGARDS Study, the logistic regression herein is based on analysis of 626 stroke events among 3,936 reported strokes and includes quite powerful predictive factors (c statistic = 0.854). As such, the additional variation introduced by this step would be minor and unlikely to substantially affect the results. We are in the midst of including this improvement to the approach and would strongly suggest that it should be included by others using multiple imputation for incident events, and we stress that this inclusion is particularly important if the number of events/reports is smaller or if there are not powerful predictors of incident events. In addition, we look forward to attempting at least to partially confirm our method using data from sources such as the National Death Index (14); however, this confirmation requires several years until the data are available and will provide information on fatal stroke events only. We also acknowledge that we have treated suspected events as independent, where multiple suspected events within individual participants are correlated. However, 2,519 of the 3,325 (76%) participants with suspected events have only a single suspected event, making the impact of this correlation relatively minor and limiting the use of hierarchical models to account for the correlations.

Advancement of statistical methods to account for missing data has been applauded as an approach to improving the conduct of epidemiologic studies (8). This approach holds promise for the following: 1) to remove biases in the estimated associations of risk factors, 2) to provide improvements in the precision of estimates through the statistical properties of the multiple imputation approach, 3) to provide substantial improvements in precision by allowing the inclusions of recently adjudicated events, and 4) to yield improved estimates of incidence rates from longitudinal cohort studies. Imputation approaches are well understood, and software is widely available, so there are few barriers to achieve these gains. In summary, we encourage the use of multiple imputation approaches for event confirmation in certain settings in longitudinal cohort studies.

ACKNOWLEDGMENTS

Author affiliations: Department of Biostatistics, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama (George Howard, Leslie A. McClure, Suzanne E. Judd, David J. Rhodes, Hemant K. Tiwari); National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland (Claudia S. Moy); Department of Medicine, School of Medicine, University of Alabama at Birmingham, Birmingham, Alabama (Monika M. Safford); Department of Medicine, School of Medicine, University of Vermont, Burlington, Vermont (Mary Cushman); Department of Neurology, School of Medicine, University of Cincinnati, Cincinnati, Ohio (Brett M. Kissela, Dawn O. Kleindorfer); and Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama (Virginia J. Howard, Paul Muntner).

This work was supported by cooperative agreement U01 NS041588 from the National Institute of Neurological Disorders and Stroke, National Institutes of Health, US Department of Health and Human Services.

The authors thank the other investigators and staff of the REGARDS Study for their valuable contributions. A full list of participating REGARDS Study investigators and institutions can be found at http://www.regardsstudy.org.

Conflict of interest: none declared.

REFERENCES

1. Kannel WB, Wolf PA, Verter J, et al. Epidemiologic assessment of the role of blood pressure in stroke. The Framingham Study. *JAMA*. 1970;214(2):301–310.
2. Wolf PA, Kannel WB, Dawber TR. Prospective investigations: the Framingham Study and the epidemiology of stroke. *Adv Neurol*. 1978;19:107–120.
3. Kannel WB, Wolf PA, McGee DL, et al. Systolic blood pressure, arterial rigidity, and risk of stroke. The Framingham Study. *JAMA*. 1981;245(12):1225–1229.
4. Wolf PA, D'Agostino RB, Belanger AJ, et al. Probability of stroke: a risk profile from the Framingham Study. *Stroke*. 1991; 22(3):312–318.
5. D'Agostino RB, Wolf PA, Belanger AJ, et al. Stroke risk profile: adjustment for antihypertensive medication. The Framingham Study. *Stroke*. 1994;25(1):40–43.
6. Seshadri S, Wolf PA, Beiser A, et al. Elevated midlife blood pressure increases stroke risk in elderly persons: the Framingham Study. *Arch Intern Med*. 2001;161(19):2343–2350.
7. Kannel WB, Wolf PA. Framingham Study insights on the hazards of elevated blood pressure. *JAMA*. 2008;300(21):2545–2547.
8. Klebanoff MA, Cole SR. Use of multiple imputation in the epidemiologic literature. *Am J Epidemiol*. 2008;168(4):355–357.
9. Howard VJ, Cushman M, Pulley L, et al. The REasons for Geographic And Racial Differences in Stroke Study: objectives and design. *Neuroepidemiology*. 2005;25(3):135–143.
10. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York, NY: John Wiley & Sons; 1987.
11. Shem S. *The House of God*. New York, NY: Dell Publishing; 1978.
12. Rosamond WD, Folsom AR, Chambless LE, et al. Stroke incidence and survival among middle-aged adults: 9-year follow-up of the Atherosclerosis Risk in Communities (ARIC) cohort. *Stroke*. 1999;30(4):736–743.
13. El-Saed A, Kuller LH, Newman AB, et al. Geographic variations in stroke incidence and mortality among older populations in four US communities. *Stroke*. 2006;37(8):1975–1979.
14. National Center for Health Statistics. *National Death Index User's Manual*. Washington, DC: National Center for Health Statistics; 1997.