A Case-by-Case Comparison of the Classification of Law Enforcement and Vital Statistics Data on Homicide

Marc Riedel
Southeastern Louisiana University

Wendy C. Regoeczi
Cleveland State University, w.regoeczi@csuohio.edu

A Case-by-Case Comparison
of the Classification of
Law Enforcement and Vital
Statistics Data on Homicide

Marc Riedel
Southeastern Louisiana University
Wendy C. Regoeczi
Cleveland State University

This study uses data from the California Linked Homicide File as a basis for evaluating the validity and reliability of homicide data. Case-by-case comparisons of variables reported by both agencies indicate that agreement between law enforcement and vital statistics data is highest with classifying homicides and victim gender and race and lowest with classifying victim age, manslaughters, and police justifiable homicides. The findings from a multilevel analysis examining what types of cases are unable to be linked over the two data-collection systems reveal that homicides involving Hispanic victims, weapons other than handguns, and family members other than intimate partners and homicides involving felonies, other nonfelonies, and negligent manslaughters have a greater likelihood of not being matched across the agencies. Death investigation systems that use medical examiners also decrease matching. The need for qualitative research examining how classification decisions are made by police and medical examiners or coroners is discussed.

Keywords: *homicide; vital statistics; Supplementary Homicide Reports; death certificates*

It is axiomatic that sound practice and policy depends on reliable and valid data about the target problem. Criminology and criminal justice have an additional problem: Their phenomenon of inquiry, crime, is socially constructed, and much of the data used in research are collected by criminal justice agencies to support their organizational goals (Riedel, 2000; Swigert, 1989).

Because official records and statistics support organizational goals, it does not follow that data meet the stringent standards of reliability and validity needed for scientific research. For example, the sociological perspective of social constructionism has repeatedly demonstrated how criminal justice statistics have been presented to promote narrow organizational agendas rather than a broadly consensual state of affairs (Best, 1990, 1999, 2001; Best & Horiuchi, 1985; Jenkins, 1994; Riedel, 1998).

The quality of homicide data is often argued to be superior to those on other crimes because of the fact that they are the most likely to be discovered, reported, and recorded. Consequently, their validity is rarely, if ever, subject to scrutiny. Homicide data are unique in that there are two nationwide reporting systems that gather detailed information on what is believed to be the same population of events. Because homicide data are reported by two agencies, it is possible to evaluate their strengths and limitations by examining the extent to which Supplementary Homicide Reports (SHR) data agrees with vital statistics (VS) data.

To date these comparisons have been limited to aggregated data no lower than the county level. The present study avoids the shortcomings of aggregated data as a unit of analysis by studying agreement using a data set that probabilistically matches SHR and VS data on a case-by-case basis. Although one previous study attempted case-by-case matching, the present study examines the amount of agreement using a recently released data set on California homicides from 1990 to 1999 with an overall probability match of 0.93 (California Department of Health Services, Epidemiology and Prevention for Injury Control Branch, Violent Injury Surveillance Program, 2001).

There are four major advantages to using case-level data. First, it is possible to make the most accurate comparisons of agreement because there are variables shared by both data sets such as age, race/ethnicity, and gender of victims.

Second, because of shared variables, evaluation of agreement can be made at different levels of aggregation. Thus, it is possible to examine agreement at the state or county level or at any of a variety of population aggregates.

Third, although both data sources report on the same event, the information gathered by the two sources overlaps rather than is identical. Although both sources gather information on victim age, SHR collects the same information on known offenders, whereas VS does not. On the other hand, VS gathers information on victim's education and marital status, whereas SHR does not. Thus, where the two data sources agree, the combined information is more valuable than what is available from a single source.

Finally, although the two data sources in California were matched with a 0.93 probability, there were 2,421 cases (out of 34,584) that were unmatched, leaving only SHR data available in these cases. Thus, it is important to explore differences between cases that were matched across the two data sources and those cases classified as SHR homicides that could not be matched with cases from the VS file.

Homicide Data

When a suspected homicide is reported, both the police (or appropriate law enforcement agency) and county medical examiners or coroners begin an investigation. The two offices have separate responsibilities, but they are generally believed to cooperate in determining whether a homicide has occurred. Medical examiners are charged with assigning a cause of death and judging whether medical evidence indicates that the death occurred by the actions of another person. Although limited to vic-

tim characteristics, medical examiner records contain information about variables such as marital status and education that are not available from police records. Police, on the other hand, have responsibilities related to criminal law. They conclude whether a criminal homicide has occurred and, if so, develop records to facilitate the investigation, arrest, and prosecution of offenders. Police records contain information about offenders and arrests and about some victim characteristics.

To satisfy the demands of national reporting programs, a subset of information is extracted from both medical examiner and police records and is forwarded to state reporting agencies. All states have VS offices to which death certificates are forwarded. Many police jurisdictions send their data to state-level agencies, but as has been discussed elsewhere, some police jurisdictions report directly to the Uniform Crime Reporting Program (UCR; Riedel, 1999).

The information collected at the state level is transmitted to national reporting programs and reported annually. Police-based data go to the UCR Program administered by the FBI, whereas copies of death certificates go to the Mortality Division of the National Center for Health Statistics. Hereafter, death certificate information will be referred to as VS records.

Unless homicide researchers obtain information directly from the official records of reporting agencies, data are taken from the SHR, one of the forms used by the UCR Program. Unlike other UCR forms, the SHR gives the ages, races/ethnicities, and genders of victims and offenders, the victim-offender relationships, the weapons used, and the locations and circumstances of each homicide.

Studies reviewed in the following section indicate that the larger the unit of analysis, the greater the amount of agreement between SHR and VS. Conversely, the smaller the units, such as states and counties, the greater the disagreement.

Previous Research

National Comparisons

In one of the earliest comparisons of UCR and VS data, Hindelang (1974) compared national UCR murder and nonnegligent manslaughter from 1940 through 1970 to those provided by the National Center for Health Statistics. Hindelang (1974) concluded that

> CHS [Center for Health Statistics] and UCR are in reasonably close agreement with respect to estimates of homicide rates in the period for which data from both sources are available; to the extent that one lends credence to the CHS results, credence is also lent to the UCR homicide results. (p. 5)

To some extent, the agreement is a function of the UCR measure used. Hindelang used annual Return A rates. Return A, crimes known to the police, is a monthly aggregate that includes offenses reported, founded, and cleared. Return A also uses estimates to account for underreporting (Riedel, 1999).

Annual Return A rates are typically higher than SHR rates. Riedel (1999) calculated Return A rates and SHR rates and compared them to VS rates from 1960 through 1994. Dividing VS rates by Return A estimates for the 32-year series resulted in a mean agreement ratio of 1.04, signifying that VS reported 4% more homicides than did Return A. However, the mean agreement ratio for VS divided by SHR revealed that VS reported 17.0% more homicides than did SHR. Overall, Return A rates agreed well with VS rates, but the difference between VS and SHR rates was substantially larger partly because of the fact that not all homicides are reported on the SHR and partly because Return A employs estimates to account for underreporting.

Rokaw, Mercy, and Smith (1990) compared SHR frequencies to VS codes (E960.0-E969.9) from 1976 to 1982. E960.0 through E969.9 are homicide codes used by the ninth edition of the International Classification of Diseases (ICD9). Murder, nonnegligent manslaughters, and justifiable homicides were included in the SHR, whereas negligent manslaughters were excluded. For the 7-year period, VS reported more homicides than did SHR. The mean difference of 1,791 more homicides for VS indicated that VS reported 9% more homicides than did SHR. Ratios of the annual total frequencies did not vary in a systematic way. Examining the proportional distributions of homicide by month showed the seasonal patterns to be very close.

Cantor and Cohen (1980) performed an extensive analysis of eight time series compiled by VS, UCR, and the Office of Management and Budget. Although there was close agreement between VS and Return A estimates from 1963 to 1973, they stress that correlations between the two time series will provide different results depending on the period chosen between 1933 and 1975.

State Comparisons

In addition to a national comparison, Hindelang (1974) also compared VS and Return A estimates for each state for 1968. Although they agree closely (the median difference value was 8), the differences were skewed toward more homicides being reported by VS.

Similarly, a state-by-state comparison of SHR and VS homicide frequencies by Rokaw et al. (1990) revealed substantially different results from national comparisons. In 11 states for the 1976 to 1982 period, VS homicide reports exceed SHR reports by more than 20%. In 4 states, SHR reports exceed VS homicide reports. The authors conclude:

> Some States had extreme variability in their annual ratios. Annual ratios [of SHR divided by VS] for New Mexico varied from 1.13 to 42.0 with a standard deviation of 15.32. Although the frequencies in many States were small for both systems, clearly the UCR system receives substantially fewer homicide reports from many states than does the mortality system. (p. 451)

In the Rokaw et al. (1990) research, California was among the states with the closest agreement between SHR and VS. For 1976 to 1982, the agreement ratio was 1.01,

indicating that VS reported 1% more homicides than did SHR (20,162 VS homicides vs. 20,015 SHR homicides).

Keppel, Weis, and LaMoria (1990) developed the Homicide Information and Tracking System to obtain information in the state of Washington. After a careful examination of information from VS, UCR, and local offices of police and medical examiners, they concluded there were 1,309 homicides in Washington from 1980 through 1986. Using this figure as a comparison base, the authors found that local medical examiners and coroners were responsible for the largest amount of under-reporting (–21.3%), followed by VS (–16.0%). The UCR underreported by 4.7%, whereas local law enforcement underreported by only 0.5%.

County Comparisons

The question of agreement between VS and SHR is particularly important when the focus of the inquiry is counties or cities within counties. If VS and SHR do not agree, then results obtained run the risk of being affected by the data source used.

Wiersema, Loftin, and McDowall (2000) examined the amount of agreement between SHR and VS records for 3,111 counties or county equivalents in the United States from 1980 to 1988. The VS data used codes E960.0 through E969.9, whereas the SHR used murders and nonnegligent manslaughters excluding negligent man-slaughters and justifiable homicides. The authors found that in 670 (22%) counties, the two homicide estimates agreed exactly. For more than two thirds of the counties (2,120) the difference was ±4 homicides or less. The full distribution of differences, however, extended from –474 to +3,121. Nine counties had differences between –474 and –30, indicating more SHR homicides were reported than were VS homicides. There were 154 counties with differences between +30 and +3,121, indicating more VS than SHR homicides were reported.

Counties reporting more VS than SHR homicides tend to contain larger cities. The counties with the largest positive differences contain Chicago, Houston, New York, Los Angeles, and Miami. For example, Cook County (Chicago) reported 8,089 VS homicides and 4,968 SHR homicides from 1980 through 1988, for a difference of +3,121. Los Angeles had 13,531 VS and 12,543 SHR homicides, for a difference of +988.

Counties with the largest negative differences, that is more SHR homicides were reported than were VS homicides, contain medium-sized cities. They were Franklin County, Ohio (Columbus); Alameda County, California (Oakland); Summit County, Ohio (Akron); San Joaquin County, California (Stockton); and Baltimore County, Maryland (Baltimore). Columbus VS reported 314 homicides and 789 SHR homicides, for a difference of –475. In California, Oakland reported 1,164 VS homicides and 1,385 SHR homicides, for a difference of –221, whereas Stockton reported 403 VS homicides and 510 SHR homicides, for a difference of –107 (Wiersema et al., 2000).

In sum, high levels of agreement between VS and UCR have employed national-level data and relied on Return A estimates. More current national comparisons of SHR, Return A, and VS also indicate a high level of agreement between VS and

Return A but a substantially lower level when VS is compared to SHR. It seems likely that early national studies showing high levels of agreement capitalized on aggregation error and the higher frequencies of Return A estimates. In addition, the research by Rokaw et al. (1990) and Wiersema et al. (2000) reveals that there are high levels of agreement among many states and counties. The extreme differences between the two data sources occur in a few states and counties.

Although not uniformly true, VS reports more homicides than do either Return A or SHR. Even when more refined units of analysis such as states and counties are used, there are more instances of higher VS than SHR frequencies than the reverse. One reason that VS generally reports more homicides than the SHR is because there is no separate category for justifiable homicides by civilians in VS, whereas these offenses have a separate classification in SHR. Thus, when murder and nonnegligent manslaughters are compared to VS homicide classifications, overreporting occurs because civilian justifiable homicides are included in the VS classification as homicides but are not included among the murders and nonnegligent manslaughters.

Many of the studies compared the SHR category of murders and nonnegligent manslaughters with E960.0 through E969.9, the most comparable homicide codes. However, Rokaw et al. (1990) compared murders, nonnegligent manslaughters, and justifiable homicides (both civilian and police) with national and state frequencies of E970.0 through E978.9, legal intervention homicides (which include police justifiable homicides). Legal intervention homicides are "injuries inflicted by the police or other law-enforcing agents, including military on duty, in the course of arresting or attempting to arrest lawbreakers, suppressing disturbances, maintaining order, and other legal action" (p. 449). Also included are legal executions. Although this was the only study to include justifiable homicides or legal intervention, the results suggest that the broader category provides the same results.

Because the focus of this article is on homicide data in California, it is worth noting its position vis-à-vis other jurisdictions. In the comparison by Rokaw et al. (1990), California reported 1% more VS than SHR homicides. Wiersema et al. (2000) found that among the five counties that reported the largest positive differences in favor of VS over SHR, three large counties reported greater differences than did Los Angeles. Stockton and Oakland were among the five cities with the largest negative differences. However, examining the trends for the two cities indicates a convergence of VS and SHR counts in the last 2 or 3 years prior to 1988.

It appears there has been only one attempt to merge SHR and VS data prior to the California Linked Homicide project. The Bureau of Justice Statistics attempted to create a comprehensive national homicide database by merging VS data (E960.0-E978.9) that included homicides and legal intervention homicides with SHR data including murders and nonnegligent manslaughters, negligent manslaughters, and justifiable homicides. Death certificates were obtained for July 1986 from all states and the District of Columbia, except Maryland, New Jersey, Virginia, and upstate New York. Of the total VS cases (1,855) and total SHR cases (1,783), 1,191 cases were matched. This left 664 unmatched death certificates and 572 unmatched SHR cases (Rand, 1993).

Method

Data

The process of merging SHR and VS is described in detail in the documentation provided by the California Department of Health Services, Epidemiology and Prevention for Injury Control Branch, Violent Injury Surveillance Program (2001). The SHR data set consists of 34,584 homicides investigated and reported to the California Criminal Justice Statistics Center from 1990 to 1999. The Department of Health Services provided the death records on a death statistical master file. Considering the goal was to link as many death records as possible to the homicide file, all 170,111 injury deaths (E800.0-E999.9) from 1990 to 1999 were used.

Integrity, formerly known as *Automatch*, performed the linkage between the two data sets. *Integrity* is a probabilistic linkage program that uses selected variables to link cases from the two data sources, assigning a final probability to the success of the linkage. Each record from the SHR file was treated independently and permitted to match with any VS record. Social security number, last name, first name, middle name, sex, age, date of homicide, date of injury, date of death, and county were used in the linkage process.[1] Including the automated and manual linking that was performed, 32,163 of the 34,584 cases were matched, for a matching rate of 93%. A total of 2,421 cases were designated homicides by law enforcement but could not be matched; thus only SHR data are available for those cases.

The analysis is limited to the years 1992 through 1998. The initial year of 1992 was made necessary by the fact that data collection on domestic violence and gang killings did not begin until 1992. This reduced the number of cases from 34,584 to 26,557. The ending point of 1998 and the ICD9 were used because the 10th ICD revision was not fully implemented in the 1999 data. Information on drive-by shootings was collected only beginning in 1996; thus we omitted these 313 cases, leaving a final data set of 26,244 cases (24,426 matched and 1,818 unmatched cases).

Variables

Individual-Level Variables

Multiple victims. A binary variable was created to denote homicides involving two or more victims.

Victim race/ethnicity. Race/ethnicity was recoded into four dummy variables: Whites, Hispanics, Blacks, and other racial/ethnic groups. Whites were the reference category.

Victim/offender relationships. We constructed the following series of relationship dummy variables: intimate partners (husband, wife, common-law husband, common-law wife, boyfriend, girlfriend, ex-husband, ex-wife, homosexual relationship); other

family (mother, father, son, daughter, brother, sister, in-law, stepfather, stepmother, stepson, stepdaughter, other family); others known to victim (neighbor, acquaintance, employee, employer, friend, other known to victim); and strangers. Strangers were the reference category.

Circumstances. Circumstances were measured by constructing the following series of dummy variables: altercations (brawls, arguments, domestic violence); felonies (robbery, rape, burglary, larceny, motor vehicle theft, arson, prostitution and commercial vice, other sex offense, abortion, narcotic drug laws, gambling, other felony); organized crime or gangs (organized crime, contract killing, contract arson); other nonfelony (child abuse, child killed by babysitter, institutional killings, sniper attack, other); negligent manslaughters (hunting and gun accidents); justifiable homicides by civilians (felon attacked police or others, resisted arrest, killed in commission of crime); and justifiable homicide by police officer (felon attacked police or civilian, attempted flight, killed in commission of crime, resisted arrest). Altercations were the reference category.

Locations. Locations were recoded into a series of dummy variables: private residence (short-term residence, victim or offender or shared or other residence); public indoor location (service station, convenience store, fast food restaurant, liquor store, other business, financial institution, warehouse, other storage area, bar, restaurant, vehicle); public outdoor locations (street, highway, park, public use area, vacant area, jail, California Youth Authority, department of corrections, school grounds); and other locations. Private residences were the reference category.

Weapons. Weapons were recoded into a series of dummy variables: handguns; other firearms (firearms, shotgun, rifle); knives; blunt objects; personal weapons (hands, feet, teeth, etc.); and other weapons. Handguns were the reference category.

County-Level Variables

Death investigation system. California uses a mixed system of coroners and medical examiners. Some of the coroner counties have sheriff-coroners. The qualifications for these positions are extremely divergent. Coroners and sheriff-coroners are elected and face minimal qualifications: "citizen of California over 18 years of age, and an elector in the county where the duties of the office are to be exercised" (Epidemiology Program Office, 2002, p. 1). In contrast, medical examiners are appointed and must be licensed physicians and surgeons with qualifications as specialists in forensic pathology. In California, there are 4 counties with medical examiners, 9 counties with coroners, 43 counties with sheriff-coroners, and 1 county with a district attorney–coroner (Trinity County) and 1 county that is a combination of both coroner and medical examiner (Los Angeles).

For the multivariate analysis, dummy variables were created to indicate counties using a (a) medical examiner and (b) sheriff-coroner. Los Angeles was classified as a

medical examiner county even though it uses a mixed system on the expectation that the existence of a medical examiner should make it comparable to other counties using medical examiners.

Population size. As a control for county size, we constructed a variable of the average population size of the county during the years under examination. We use the log of the variable to adjust for its skewed distribution.

Analysis

The analysis was carried out in two stages. The first focused on the amount of agreement between shared variables for the two data sources. Comparisons were made for crime status, gender, age, and race/ethnicity. We also compared these variables by county.

The second analysis focused on the differences between the 24,426 matched cases and the 1,818 unmatched cases. For this analysis, the nested structure of the data required the use of a program that could incorporate the multilevel nature of the data. These analyses were carried out using the hierarchical linear modeling (HLM) software developed by Raudenbush, Bryk, and Congdon (2001). By using HLM, the complex error structure of the data is taken into account. The assumption under ordinary least squares regression is that the residuals are independent. However with nested data, all homicide cases in a particular county have the same error, creating a correlation. Under such circumstances, estimation by iterative maximum likelihood is appropriate. Because the outcome variable in this analysis is binary ($Y = 1$ if the cases were not matched, $Y = 0$ if they were matched), violating the assumption of normality at level 1, we use an approach referred to as the hierarchical generalized linear model (HGLM; see Raudenbush, Bryk, Cheong, & Congdon, 2000, for a detailed discussion of when and how to use HGLM models).

In the multivariate analysis, we examine the probability that cases will be unmatched across the two agencies on the basis of the following victim and offense attributes: multiple victims, race/ethnicity of the victim, victim-offender relationship, circumstances surrounding the offense, location, and weapon.

Hypotheses Regarding Effect of Death Investigation System

An additional factor hypothesized to have an effect on the likelihood of a match was the type of death investigation system in place in the county in which the death occurred. We hypothesize that the type of death investigation system will have the following effects on the probability of matching cases. First, we expect that the mixed nature of the system and the predominance of positions having minimal qualifications (i.e., coroners and sheriff-coroners) will result in variation across counties in the likelihood of successfully matching cases. Second, we hypothesize that some of this variation can be explained by county size, hence the need to control for the total population of the county. Third, we also hypothesize that the presence of a medical examiner

versus other death investigation systems will have an effect on the probability of matching cases. Given their medical (as opposed to law enforcement) backgrounds and the nonelected nature of their positions, medical examiners may function as more independent decision makers regarding causes of death than may coroners or sheriff-coroners. Thus, we expect more disagreement among decisions made by police and medical examiners than coroners, thus increasing the probability that cases will not be matched in counties using medical examiners.

Results

Agreement Between Common SHR and VS Variables

Figure 1 gives the relationship among SHR and VS variables for the two data sets where matching was possible. For all cases, A ∩ B indicates variables reported by both data sources.

The list of SHR variables in Figure 1 is no doubt familiar to many homicide researchers. Less familiar to SHR users are the VS variables on the right side of Figure 1. Researchers interested in homicide events will find variables such as victim education useful as indicators of social class, measures consistently absent from law enforcement data.
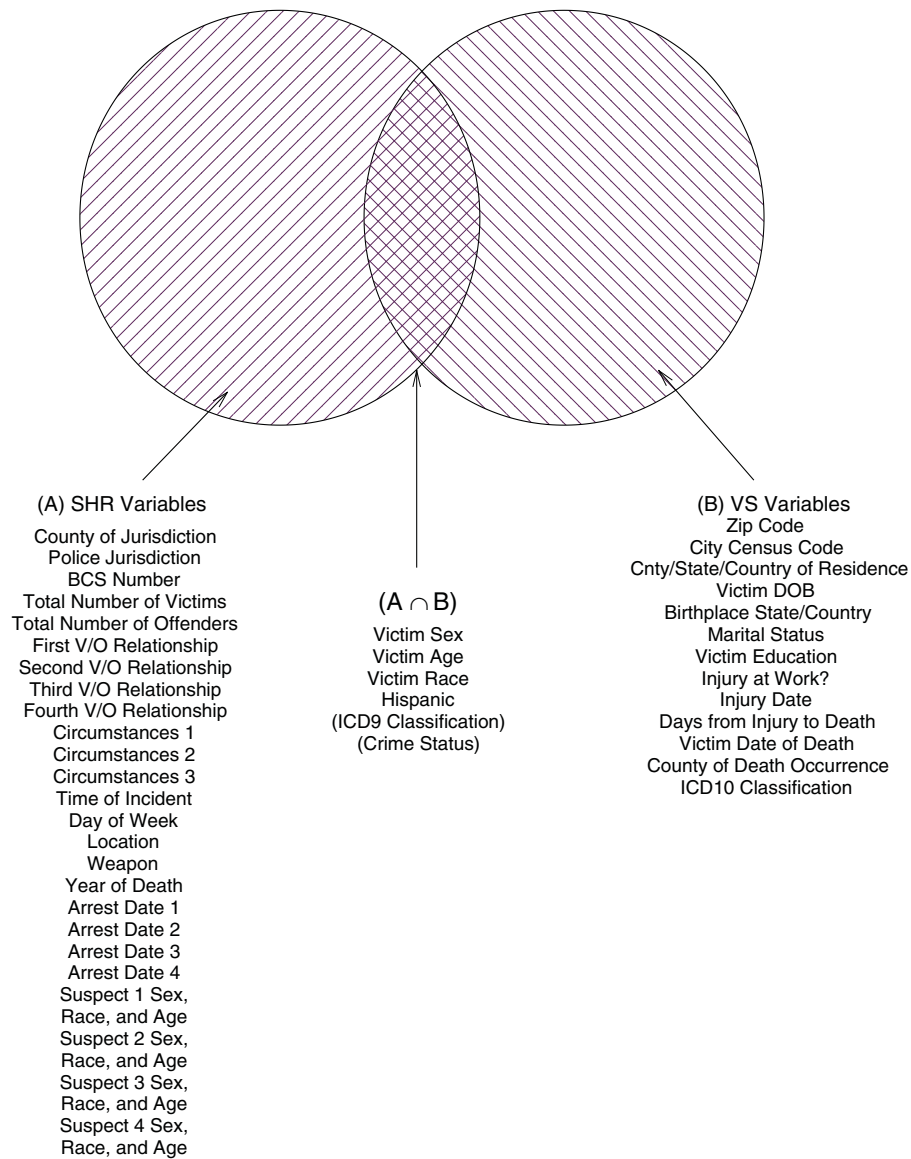
For variables reported by both sources, victim sex and age lend themselves to easy comparison. For the SHR, Hispanics are represented by one code under the race variable, whereas for VS, Hispanic is a separate variable comprised of several codes for different Hispanic ethnic groups and whether they were born in the United States. The latter codes were collapsed to make them comparable to the SHR race variable to study agreement. The ICD9 and crime status are comparable in a different way and are discussed in the following section.

Crime Status

The crime status variable used by the SHR includes willful homicides, police and civilian justifiable homicides, and manslaughters. To compare SHR codes with VS, we collapsed the ICD9 classification into homicides, legal interventions, and other kinds of personal injury.

The results show that VS and SHR generally agree (98.1%) on the broad category of homicides used by ICD and willful homicides or murder used in California (see Table 1). Of events classified as manslaughters by the SHR, 66.3% are classified as homicides by VS, whereas 33.2% are classified in other categories such as accidents. In attempting to conceptually match SHR and VS classifications, Rokaw et al. (1990) suggest that negligent manslaughters are most comparable to accidents or homicides with undetermined intent. Table 1 indicates that a substantial number (66.3%) are also classified as homicides by VS; about one third are classified as other types of personal injuries, which may include accidents or undetermined intent.

Figure 1
Relationship Among SHR and VS Variables: The Matched File

(A) SHR Variables

County of Jurisdiction
Police Jurisdiction
BCS Number
Total Number of Victims
Total Number of Offenders
First V/O Relationship
Second V/O Relationship
Third V/O Relationship
Fourth V/O Relationship
Circumstances 1
Circumstances 2
Circumstances 3
Time of Incident
Day of Week
Location
Weapon
Year of Death
Arrest Date 1
Arrest Date 2
Arrest Date 3
Arrest Date 4
Suspect 1 Sex,
Race, and Age
Suspect 2 Sex,
Race, and Age
Suspect 3 Sex,
Race, and Age
Suspect 4 Sex,
Race, and Age

$(A \cap B)$

Victim Sex
Victim Age
Victim Race
Hispanic
(ICD9 Classification)
(Crime Status)

(B) VS Variables
Zip Code
City Census Code
Cnty/State/Country of Residence
Victim DOB
Birthplace State/Country
Marital Status
Victim Education
Injury at Work?
Injury Date
Days from Injury to Death
Victim Date of Death
County of Death Occurrence
ICD10 Classification

Note: SHR = Supplementary Homicide Reports; VS = vital statistics.

VS has no separate classification for civilian justifiable homicides, thus it is not surprising that 96.2% of them are classified as homicides. Among the 476 cases classified as civilian justifiable homicides, only 2.7% are classified as legal intervention, and less than 1.0% are classified by VS as all other causes of injury.

We also examined the agreement among data sources for California counties. We focused only on 20 of 57 California counties with 100 matched homicides or more from 1992 through 1998. (No homicides were recorded for Alpine County.) The percentage of events classified as homicides by SHR and VS varied little throughout the selected counties; the lowest was for Santa Clara County (95.4%).

The numbers of manslaughter cases for each of the 20 counties were so small as to preclude any meaningful comparisons. Even for the county with the largest number of homicides, Los Angeles (9,819), there were only 45 SHR manslaughters, 43 of which were classified as homicides by VS.

Table 1 indicates that SHR classifies more homicides as justified killings by police than does VS. Of the 823 homicides classified as justifiable by police in the SHR, VS classified only 44.1% of these cases as legal intervention homicides. Sherman and Langworthy (1979) also found that police classify more homicides by police as justifiable than do medical examiners. It is tempting to suggest that police may classify more police shootings as justifiable to prevent police officers from being charged with an offense that carries criminal liability.

However, a different conclusion is supported when the 20 counties with more than 100 homicides are examined. In 10 counties, for homicides classified in the SHR as police justifiable homicides, VS classified more of them as legal intervention than homicides. For example, in San Bernardino County, 63 of 1,408 homicides were classified as justifiable homicides by police. Of those, VS classified 79.4% as legal interventions and 20.6% as homicides. In 7 counties, the percentages were reversed, with police justifiable homicides more frequently reported by VS as homicides than legal intervention. Inspection suggests these are counties with larger numbers of homicides. For example, in Los Angeles County, 302 of 10,366 homicides were reported as justifiable homicides by police. Of those, an impressive 90.4% were reported by VS as homicides and only 9.3% were reported as legal intervention.

In one county (Fresno), SHR and VS agreed for 10 events, and another 10 events were classified as justifiable homicide by SHR and as homicides by VS. There were two counties (Merced and Monterey) with a very small number of homicides that were classified as police justifiable by SHR and legal intervention by VS.

The results for Los Angeles County help to explain why more police justifiable homicides were reported as VS homicides in Table 1. Of the 442 police justifiable homicides that were reported by VS as homicides, 61.8% of them came from Los Angeles County. Thus, before concluding that police report more justifiable homicides than medical examiners, it is important to examine variations by counties. The result may be a function of a few counties with very large numbers of homicides. Additional support for this point of view comes from other studies, some of which used different units of analysis that show legal intervention homicides mixed with other homicides (Lapidus, Gregorio, & Hansen, 1990; Moyer, Boyle, & Pollock,

Table 1
ICD by SHR Crime Status Classification, 1992 to 1998

	SHR Crime Status (%)				
ICD	Willful Homicides	Manslaughter	Civilian Justifiable	Police Justifiable	Total
Homicides	98.1	66.3	96.2	53.7	21,670
Legal interventions	0.0	0.5	2.7	44.1	382
All other	1.8	33.2	1.1	2.1	471
Total	21,034	190	476	823	22,523

Note: ICD = International Classification of Diseases; SHR = Supplementary Homicide Reports.

1989; Van Court & Trent, 2004; Wiersema et al., 2000). Of course, this does not answer the larger question of why there is variation across counties.

Gender

A comparison of the amount of agreement between the SHR gender variable and the VS classification of gender is provided in Table 2. Not unexpectedly, the agreement on victim gender classification for VS and SHR is very close. Among 20,243 SHR male victims, only 0.2% were classified as female by VS; among 4,183 female victims, only 0.7% were classified as male by VS.

We made separate cross-tabulations for murders, manslaughters, and justifiable homicides and found similar small amounts of misclassification. Eight counties either agreed with the results in Table 2 or had a 1.0% or less disagreement. The remaining counties showed complete agreement between SHR and VS classifications.

Age

The mean SHR age was 30.04 (δ = 14.77), whereas the VS mean age was 30.19 (δ = 14.68). The correlation between the SHR and VS age variables was .972. There were 13 VS cases missing information on age. A more detailed examination of differences between ages reported in the SHR and VS is given in Table 3.

Table 3 indicates a large spread of differences regardless of whether VS ages are greater than SHR ages or the reverse. For example, there were three cases where the VS age was between 70 and 90 years greater than the SHR age. Fortunately, these extreme cases were very few: Including 2 years on either side of identical ages (79.9%) captured 95% of all the cases.

We also examined age differences by counties. Examining identical age percentages for the 20 counties with 100 or more homicides reveals 10 counties were greater than 79.9% with the highest percentage of agreement (88.1%) occurring in Merced County. Eight counties had a percentage agreement less than 79.5%, with the lowest agreement (59.1%) being San Francisco County. Los Angeles County had an agreement percentage of 79.5%. There is a tendency for counties with fewer than 100 homi-

Table 2
ICD by SHR Victim Gender Classification

	SHR Victim Gender (%)		
ICD Victim Gender	Male	Female	Total
Male	99.8	0.7	20,223
Female	0.2	99.3	4,203
Total	20,243	4,183	24,426

Note: ICD = International Classification of Diseases; SHR = Supplementary Homicide Reports.

Table 3
Differences in Ages Between SHR and VS

Age Difference (in years)	Frequency	%
SHR greater than VS		
50-69	2	0.0
30-49	6	0.0
10-29	126	0.5
3-9	371	1.5
1-2	2,407	9.9
Identical agreement		
0	19,501	79.9
VS greater than SHR		
1-2	1,330	5.4
3-9	363	1.5
10-29	223	0.9
30-49	65	0.3
50-69	16	0.1
70-90	3	0.0
Total	24,413	100.0

Note: SHR = Supplementary Homicide Reports; VS = vital statistics.

cides to have a greater amount of agreement, probably reflecting fewer opportunities for disagreement.

Race/Ethnicity

To make the SHR and VS race variables comparable, we collapsed the 21 categories used by VS into 17 categories to match the SHR classification. Unlike the SHR, which includes Hispanics as part of its race classification, VS has a separate variable that distinguishes a variety of Hispanics and their place of origin. To make two comparable variables, we collapsed the Hispanic variable into one category and merged it with the VS race variable.[2]

Table 4 examines the amount of agreement between the two data sources on race/ethnicity when VS Hispanics are defined to include any racial group. For the three

largest racial/ethnic groups (White, Hispanic, and Black), the agreement between SHR and VS classifications is high, ranging from 93.2% for Whites to 97.1% for Blacks. It is unclear what factors determine the fluctuations in classification of racial/ethnic groups other than Whites, Hispanics, and Blacks. For example, the two data sources agree on the classification of Japanese (80.6%), even though there were only 36 victims. However, another group almost the same size, Pacific Islanders ($n = 30$), has an agreement of only 26.7% for the two data sources.

To simplify more detailed comparisons, we recoded SHR and VS racial/ethnic variables into White, Hispanic, Black, and Other. For the collapsed *Other* category, the percentage agreement between the two data sources was 84.8%.

Of the 20 counties with more than 100 homicides, 18 had percentage agreements for White victims that ranged between 92.2% and 97.3%. Monterey County had complete agreement for 53 White victims. At the lower end of agreement, San Francisco County ($n = 210$) and Solano County ($n = 75$) had White agreement percentages of only 71.4% and 86.7%, respectively. The general percentage agreement in Table 4 was identical for Whites and Hispanics (93.2%). The range of agreement for the 20 counties for Hispanics was between 90.0% and 97.9%. San Francisco County had a percentage agreement for Hispanics for the two data sources of 82.4%. For 20 counties, the percent of agreement for Blacks ranged from 85.0% to 100.0%. The range of agreement among the counties for the *Other* category was from 73.8% to 91.3%. One county, Kern, had a percentage agreement of 60.0%.

When considering racial/ethnic groups other than Whites and Hispanics, percentage agreements should be viewed with caution. Not only is there a difference between the volume of homicides for Whites and Hispanics, minorities are distributed differently among the counties. For example, out of 500 homicides in Kern County, only 15 were homicides involving racial/ethnic groups other than Black, White, or Hispanic.

Comparing Matched and Unmatched Cases

The second part of the analysis examines factors that influence the likelihood of an SHR case not being matched with a VS case. We began by examining an initial model, akin to a one-way ANOVA, that did not include any individual-level or county-level predictors. This model produced an estimate of the variance in the dependent variable (unmatched vs. matched) at the individual and county levels. It also permitted an assessment of whether any significant differences in matching exist across counties (see Bryk & Raudenbush, 1992). The estimated intercept term (–2.867) produced a predicted mean, county-level unmatched rate of 5.7%. The variance of the intercept term was estimated to be 0.104, which was highly significant ($p < .001$), indicating that the probability of cases not matching varied considerably across counties.

In the next model we added the individual-level predictors. The slope coefficients for all predictors were treated as fixed, with one exception: We included a random effect for police justifiable homicides as a test of whether there is significant variation in the effect of this variable on the likelihood of cases not matching across the different counties. The variances of both the intercept and the coefficient for police justifiable

Table 4
Percentage Agreement Between VS Race/Ethnicity and SHR Race/Ethnicity

Victim Race/Ethnicity	Percentage Agreement	Number of Cases
Non-Hispanic White	93.2	5,573
Hispanic	93.2	10,395
Black	97.1	6,848
American Indian	45.7	116
Chinese	57.6	85
Japanese	80.6	36
Filipino	84.2	196
Other	5.0	380
Pacific Islander	26.7	30
Asian	14.1	411
Cambodian	19.3	29
Guamanian	66.7	3
Korean	83.3	96
Laotian	73.9	23
Samoan	77.8	36
Vietnamese	86.8	152
Hawaiian	0.0	1

Note: VS = vital statistics; SHR = Supplementary Homicide Reports.

homicides were significant. Therefore we ran a final model in which we introduced county-level variables as predictors of these between-county differences. In particular, we included a measure of whether the county uses a medical examiner as a predictor of variation in the intercept, as well as population size as a control, and the use of a sheriff-coroner as a predictor of variation in the slope for police justifiable homicides. The results of this final model are presented in Table 5.[3]

Examining first the individual-level predictors, we find that compared to White victims, the odds of no match increased by a factor of 1.667 for Hispanic victims.

The only significant coefficient for the effect of victim-offender relationship occurs with respect to other family members, which have a greater likelihood of not being matched compared to cases involving strangers.

In terms of the circumstances surrounding the offense, compared to altercations, the odds of no match increased by a factor of 1.451 for felonies, 1.910 for other nonfelonies, and 1.781 for negligent manslaughters. Compared to altercations, cases involving gangs and organized crime have decreased odds of being unmatched. These latter types of cases may fit so strongly with conceptions of intentional killings that death certificates are routinely available in injury death files, making the likelihood of finding a match higher.

Compared to homicides committed with a handgun, the odds of no match increase by factors of 1.600, 3.343, 3.785, and 5.043 for knives, blunt objects, personal weapons, and other weapons, respectively.

By themselves, homicides classified by police as justifiable are no more or less likely to be matched. However, we find a significant interaction between police justifi-

Table 5
Hierarchical Logistic Regression Analysis
of Unmatched Versus Matched Homicides

Variable	B	SE	Odds Ratio
Intercept	–3.923***	.208	—
Multiple victims	0.123	.117	1.131
Race			
Hispanic	0.511***	.086	1.667
Black	–0.083	.104	0.920
Other race	0.246	.140	1.279
Relation			
Intimate partners	0.066	.145	1.068
Other family	0.297*	.131	1.346
Other—Known to victim	0.092	.086	1.096
Circumstances			
Felonies	0.372***	.094	1.451
Organized crime or gangs	–0.355**	.120	0.701
Other nonfelonies	0.647***	.114	1.910
Negligent manslaughters	0.577*	.235	1.781
Civilian justifiable	0.369	.214	1.446
Police justifiable	–0.682	.438	0.506
Police Justifiable × Sheriff or Coroner	1.279*	.504	3.593
Location			
Public indoor	–0.025	.111	0.975
Public outdoor	0.150	.085	1.162
Other location	0.510	.325	1.665
Weapons			
Firearms	–0.215	.154	0.807
Knives	0.470***	.108	1.600
Blunt objects	1.207***	.130	3.343
Personal weapons	1.331**	.118	3.785
Other weapons	1.618***	.109	5.043
Medical examiner	0.401*	.191	1.493
Total population (logged)	–0.305*	.140	0.737

Note: N = 18,855. Reference categories are White victims, strangers, altercations, private indoors, and handguns.
*p < .05. **p < .01. ***p < .001.

able homicides and counties using sheriff-coroners. In particular, the odds of no match increase for police justifiable homicides in sheriff-coroner counties. Sheriff-coroners may be more likely to classify the death as something other than a homicide, resulting in a death certificate not used in the linkage process, given their double duty as both a death investigator and law enforcement officer.

Finally, we found support for the hypothesis that counties using medical examiners will have a greater likelihood of unmatched cases because of more disagreement surrounding decisions made by police and medical examiners than by coroners. The odds

of the case not being matched are about 1.5 times greater in counties with medical examiners.

Discussion

It is important to keep in mind the simple fact that what characterizes the unmatched cases is the inability to match them. The difficulty is that we can only speculate about the unmatched cases because these cases have only SHR data. Rand's (1993) cogent conclusion is appropriate here:

> Differences between cases in the files are to a great degree the result of differences in the two programs' purposes and procedures. Basically, the UCR measures crimes, of which death is one outcome. The Mortality System measures deaths, of which crime is one cause. (p. 112)

Felonies, nonfelonies, negligent manslaughters, and justifiable homicides by civilians are legal classifications and have limited or no parallels in ICD classifications, which pose difficulties in attempting to match the two data sources.

In the case of felonies, the increased probability of not matching may reflect the nature of these types of homicides, which more often involve individuals with no prior relationship and may be less likely to occur in the presence of witnesses than homicides stemming from altercations. Under these conditions, it may be easier to dispose of the body, and matching will be made more difficult because either the death certificate will lack sufficient information (because of decomposition of the body before it is found) or will not exist (because a body must be found for a death certificate to be filed but not for an SHR to be completed).

Nonfelonies may have greater odds of not matching because this category includes child abuse cases that may have been assigned a natural cause of death. The increased probability of not matching for negligent manslaughters may reflect a situation where these incidents are classified as natural causes of death by medical examiners and coroners and therefore would have been excluded during the linkage, which was limited to injury deaths.

One possible explanation for the greater likelihood a case will be unmatched when the victim is Hispanic is that some of these victims may have entered the United States illegally, making it difficult to obtain complete information for the death certificate without the assistance of a social security number or next of kin.

The increased probability that cases involving other family members will be unmatched likely reflects a situation where death certificates for infant or child homicide victims are not given an E-code. Although all injury deaths were included in the linkage process, a recent study indicates that a number of child abuse cases appear to have been coded on death certificates as naturally occurring (e.g., sudden infant death syndrome), leading them to be assigned a natural cause of death (Van Court & Trent, 2004).

Homicides committed with any type of weapon other than a firearm increases the odds that the case will be unmatched. It is possible that homicides committed with a gun produce stronger evidence of an injury death. In other words, when one person kills another using only his or her own hands or feet, by suffocation, poisoning, and so on, the question of whether the death was the result of a personal injury may be more open to dispute, leading some to be classified as natural and thus excluded from the linkage process. Furthermore, the noise resulting from gunshots may increase the likelihood of the discovery of the body early on in the investigative process, leading to more complete information on the death certificate and thereby increasing the chances of a match. Both of these explanations are supported by the finding that other firearms do not significantly increase or decrease the odds of matching compared to handguns, thus suggesting it is something about the entire category of guns that makes a match more likely.

Several conclusions can be drawn from this research. First, both data sources agree very well as to whether the event should be called a homicide. The agreement from different comparisons is approximately 98%.

Second, the view by Rokaw et al. (1990) that negligent manslaughters are most similar to accidents or homicides of undetermined causes needs to be examined further; this research found that almost two thirds were classified as homicides by VS. Third, data on justifiable homicides by police have to be used with caution. The current research indicates the SHR has a tendency to report more justifiable homicides than VS, but this is influenced by county frequencies. The underreporting by VS seems to characterize larger counties.

Fourth, agreement between the two data sources was excellent for gender and primarily for Whites, Blacks, and Hispanics. Other racial/ethnic groups had lower percentages of agreement. The two data sources agreed exactly on age of the victim in only about 79% of the cases, but including 2 years on either side of the identical age category increased the agreement to 95%.

Because the California Linked Homicide File is currently the only data set of its kind, there are no directly comparable studies on the validity and reliability of homicide classifications in other states. That said, there is little reason to expect that California would be unique in terms of some of the obstacles impeding accurate reporting of homicides through the SHR and VS systems. For example, the underreporting of child abuse cases in VS is likely to be an issue across all states. The politically sensitive nature of classifying homicides by police may well be a factor across all law enforcement jurisdictions in the country, leading to disagreements in SHR and VS reports. California is also not alone in experiencing variations across counties in the type of death investigation system. There are 17 other states, including Illinois, Minnesota, New York, Pennsylvania, Texas, and Washington, using mixed medical examiner and coroner or sheriff-coroner systems. Thus, variation in the lack of agreement in classifying homicides is probable across counties in other geographic locations as well.

Several important policy implications result from the findings of this study. In particular, policy initiatives aimed at reducing levels of lethal violence often develop in response to perceived or real increases in certain forms of homicide. Although the cur-

rent research indicates relatively high levels of agreement in the classification of SHR willful homicides and VS homicides, the extent to which classifications of other subtypes of homicide agree raises some concerns. For example, the significant degree of disagreement over the classification of police justifiable homicides implies that whether a particular city is perceived as having an unacceptably high level of police shootings necessitating policy intervention may very well depend on which data source is used to study the phenomenon. Furthermore, although there were high levels of agreement across the two data sources for victim age and gender, agreement in the classification of many racial and ethnic groups, including Native Americans, was low. Consequently, researchers and policy makers may formulate different solutions for addressing the nature and extent of homicide among the Native American population depending on from which data source this information is derived.

Finally, there are some general recommendations with respect to using the California Linked Homicide File for research. As mentioned earlier, the agreement between the two data sets on a variety of measures is in excess of 95%. Although this amount of agreement is high, it seems appropriate that statistical comparison should set an alpha level of .01, rather than the customary .05, to avoid Type I error. A second caution is that researchers should examine county comparisons carefully. In every comparison done here, there were always a few counties that had aberrant results. As the Wiersema et al. (2000) research indicates, county-level data show variations that are cancelled out in larger units of analysis.

This research suggests the next step might be qualitative research. There is a need to examine how classification decisions are made by police and medical examiners. What we suggest as a future step in research is the selection of one or two counties in the same state in which there is high agreement and one or two counties in which there is little agreement between the two sources. After selecting these two sets, what is needed is a detailed examination of how law enforcement and medical examiners in these jurisdictions make classification decisions, of how they compile reports, of how they interact with each other particularly when they disagree as to whether a homicide occurred, and of the extent of their reporting. In short, what are the differences in processing between jurisdictions in which medical and law enforcement decisions agree and jurisdictions in which they disagree? These findings could ultimately form the basis of a more uniform set of procedures to be implemented across jurisdictions and ultimately bring the classification of SHR and VS reports more in-line with each other.

Perhaps the most general conclusion has been given elsewhere (Riedel, 1999). There is a

> relative absence of information about the validity of the initial classifications in either sources. There is little encouraging research about the accuracy of medical classification of homicide and none about the accuracy of police decisions. (p. 93)

We believe the current research serves as an important first step toward filling this gap, and we hope it will serve as an impetus for further efforts in assessing the taken-for-granted quality of homicide data.

Notes

1. Five blocking passes provided a means to narrow down the number of potential matches: (a) pass 1—social security number (SSN) and sex; (b) pass 2—New York State Identification and Intelligence System (NYSIIS) of last name and NYSIIS of first name; (c) pass 3—year and month of incident (homicide file) and year and month of death (vital statistics, VS, file) and sex; (d) pass 4—year and month of incident (homicide file) and year and month of injury (VS file) and sex; (e) pass 5—NYSIIS of last name with NYSIIS of first name, NYSIIS of last name with NYSIIS of middle name, county of homicide, and county of death. Pairs of records that met the blocking criteria were assigned a weight based on the matching variables employed in the pass. Matches were then studied to determine whether they met one of the following criteria: (a) exact match on SSN and two out of three name matches; (b) two out of three names and date of injury or death; (c) two out of three names and either date of injury or death within 10 days, or county codes match or are contiguous and age within 10 years; (d) exact match on date of injury or death and first or last name match, and county codes match or are contiguous or age within 5 years.

2. Of 10,097 Hispanic victims, 85% were Mexican American, and 11.7% were from Central or South America. The remainder were from Puerto Rico or Cuba or were classified as other Spanish or Hispanic.

3. Results for the unit-specific model are presented.

References

Best, J. (1990). *Threatened children: Rhetoric and concern about child-victims*. Chicago: University of Chicago Press.

Best, J. (1999). *Random violence: How we talk about new crimes and new victims*. Berkeley: University of California Press.

Best, J. (2001). *Damned lies and statistics*. Berkeley: University of California Press.

Best, J., & Horiuchi, G. T. (1985). The razor blade in the apple: The social construction of urban legends. *Social Problems*, *32*, 488-499.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models*. Newbury Park, CA: Sage.

California Department of Health Services, Epidemiology and Prevention for Injury Control Branch, Violent Injury Surveillance Program. (2001). *Linked homicide file, 1990-1991*. Sacramento, CA: Author.

Cantor, D., & Cohen, L. E. (1980). Comparing measures of homicide trends: Methodological and substantive differences in the vital statistics and Uniform Crime Report time series (1933-1975). *Social Science Research*, *9*, 121-145.

Epidemiology Program Office. (2002). *Death investigation system descriptions*. Retrieved February 25, 2002, from http://www.cdc.gov/epo/dphsi/mecisp/california.htm

Hindelang, M. J. (1974). The Uniform Crime Reports revisited. *Journal of Criminal Justice*, *2*, 1-17.

Jenkins, P. (1994). *Using murder: The social construction of serial homicide*. New York: Aldine.

Keppel, R. D., Weis, J. G., & LaMoria, R. (1990). *Improving the investigation of murder: The homicide information and tracking system (HITS)* (National Institute of Justice Final Report No. #87-IJ-CX-0026). Washington, DC: Government Printing Office.

Lapidus, G. D., Gregorio, D. I., & Hansen, H. (1990). Misclassification of childhood homicide on death certificates. *American Journal of Public Health*, *80*, 213-214.

Moyer, L. A., Boyle, C. A., & Pollock, D. A. (1989). Validity of death certificates for injury-related causes of death. *American Journal of Epidemiology, 130*, 1024-1032.

Rand, M. R. (1993). The study of homicide caseflow: Creating a comprehensive homicide dataset. In *Questions and answers in lethal and non-lethal violence: Proceedings of the Second Annual Workshop of the Homicide Research Working Group* (pp. 103-118). Washington, DC: Government Printing Office.

Raudenbush, S., Bryk, A., Cheong, Y. F., & Congdon, R. (2000). *HLM 5: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.

Raudenbush, S., Bryk, A., & Congdon, R. (2001). *HLM 5: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.

Riedel, M. (1998). Counting stranger homicides: A case study of statistical prestidigitation. *Homicide Studies*, *2*, 206-219.

Riedel, M. (1999). Sources of homicide data: A review and comparison. In M. D. Smith & M. A. Zahn (Eds.), *Homicide: A sourcebook of social research* (pp. 75-95). Thousand Oaks, CA: Sage.

Riedel, M. (2000). *Research strategies for secondary data: A perspective for criminology and criminal justice*. Thousand Oaks, CA: Sage.

Rokaw, W. M., Mercy, J. A., & Smith, J. C. (1990). Comparing death certificate data with FBI crime reporting statistics on U.S. homicides. *Public Health Reports*, *105*, 447-455.

Sherman, L. W., & Langworthy, R. H. (1979). Measuring homicide by police officers. *Journal of Criminal Law and Criminology*, *70*, 546-560.

Swigert, V. (1989). The discipline as data: Resolving the theoretical crisis in criminology. In S. F. Messner, M. D. Krohn, & A. Liska (Eds.), *Theoretical integration in the study of deviance and crime: Problems and prospects* (pp. 129-135). Albany: State University of New York Press.

Van Court, J., & Trent, R. B. (2004). Why didn't we get them all? Analyzing unlinked records in California's Linked Homicide File. *Homicide Studies*, *8*, 311-321.

Wiersema, B., Loftin, C., & McDowall, D. (2000). A comparison of Supplementary Homicide Reports and national vital statistics system homicides estimates for U.S. counties. *Homicide Studies*, *4*, 317-340.