

# Quality of Chicago Supplementary Homicide Reports Data Compared to the Chicago Homicide Dataset

Carolyn Rebecca Block, Thomas D. Patterson and Daniel Dick  
Illinois Criminal Justice Information Authority  
September 13, 1999

The Supplementary Homicide Reports (SHR) of the Uniform Crime Reports has been for many years a foundation of homicide research in the United States (Maltz, 1999). The SHR is easily accessible, available in the National Archive of Criminal Justice Data (Fox, 1996), and contains more detail about each incident than available in the Uniform Crime Reports for other violent crimes. Policy decision-makers depend upon it to answer such basic questions as the trends in homicide rates of young offenders (Fox, 1997), the effectiveness of gun-carry laws (Lott & Mustard, 1997), the risk of death due to child abuse (Maltz, 1998), the number of domestic violence homicides (Langford, *et al.*, 1998), or the proportion of homicides committed by a stranger (Riedel, 1998; Williams & Flewelling, 1987).

Since the SHR has been and continues to be the source of much of our knowledge about lethal violence in the United States, it is important to know just how good that source is. Is the SHR complete enough and accurate enough to be a basis of policy decisions? To evaluate the completeness and accuracy of SHR data in one city, the Bureau of Justice Statistics (BJS) asked the Illinois Criminal Justice Information Authority to conduct a case-by-case comparison of the 1,784 homicides recorded in the Chicago Homicide Dataset (CHD) in 1993 or 1994 to homicides recorded in the SHR for those years. BJS wanted to know whether SHR data on offender characteristics was systematically different from data in police files, and if so, to describe those differences. The results provide information about the quality of SHR data in Chicago, and by extension, about SHR quality in other cities with similar situations.

## Methodology

### Case-by-Case Completeness and Accuracy

Maltz (1999:35) defines complete as the degree to which the SHR contains all of the homicide cases contained in some comparison data file or set of data files. Data sets used as a standard of comparison have included data collected by the National Center for Health Statistics (Riedel, 1990, 1993; 1999; Rand, 1993; Rokaw, *et al.*, 1990), monthly Uniform Crime Reports totals (Snyder, 1996:10-11; Bailey & Peterson, 1995: 184), detailed police records (Maxfield, 1989), and public health or court records (Langford, *et al.*, 1998). However, with two exceptions (Rand, 1993 and Langford, *et al.*, 1998), these comparisons have been based on monthly or yearly totals. The results of both of the exceptions demonstrate the danger of relying on totals as an indicator of completeness. Like them, the Chicago analysis was based a case-by-case comparison. We attempted to find the matching homicide in the SHR for each homicide in the CHD that was booked in 1993 or 1994.

A homicide is not necessarily entered into police records (booked) as a homicide on the same day that the incident occurs. The victim may die the next day or even years after the attack that caused the death. In addition, the police or medical

examiner's investigation that determines that the death was a homicide may take a long time. Arson murders, for example, are often booked as a homicide much later than the date of death. However, it is not until the homicide is booked that a report is sent to the SHR. The SHR data contain one date, which is unlabeled but appears to be the booking date. For a fair comparison, therefore, we selected from the CHD those homicide cases that had been *booked* in 1993 or 1994. It was these 1,784 CHD cases that we tried to find in the SHR. This is in contrast to Rand (1993) and Langford, *et al.* (1998), who used the date of death as the basis of comparison.

Rand's (1993) analysis was nationwide for one month, July, 1986, and was based on a case-by-case comparison of SHR records to death certificates. There were 1,191 cases that matched, which accounted for 67% of the 1,783 SHR cases and 64% of the 1,855 death certificates. Rand found three main reasons for the 664 "homicides" found in death certificates but missing in the SHR for that month. In some cases, there were errors in the death certificates. A homicide might have been given the code for a suicide, or the date of death might have been incorrectly recorded. Second, Rand's criteria for case-by-case matching included an exact match on victim's age. A difference of one or two years in the age estimated by the two sources may have accounted for many of the 664 cases. Third, the date of death in some of the 664 death certificates may have been different from the SHR "booking" date.

Langford, *et al.* (1998) focused on intimate partner homicides in Massachusetts over five years (1991 through 1995). They compiled a multiple-source comparison database from death certificates, news articles, district attorney records, and reports from domestic violence advocacy agencies, including Massachusetts residents killed in adjacent states (which are not included in the SHR) and excluding residents of other states killed in Massachusetts (which are included in the SHR). They defined intimate partner broadly as any homicide in which the motive for the homicide was a "dispute" between intimate partners or ex-intimate partners, included third parties killed as a second or only victim, for example children, as well as killings of a sexual rival. The matching criteria were less stringent than Rand's (1993) had been, but probably just as accurate. For example, 47% of the unmatched fields in matched cases represented a difference of one year in the age recorded in the two sources. The total comparison database of victims who died in Massachusetts contained 190 cases, 147 intimate partner victims and 43 other victims. Of the 147 partner victims in the comparison database, 32 (22%) had no match in the SHR, and an additional 25 (17%) were in the SHR but not recorded as an intimate relationship.

In the Chicago study, we were not only concerned about missing cases at the victim level, but also about missing cases at the offender level. To be complete, a SHR offender-level database should include an offender record for each offender in the CHD. To the degree that it does not, and to the degree that the missing offender records are systematically different from the offender records that are present in the SHR, the SHR data would be misleading for policy and research decisions.

Accuracy, in the Chicago analysis, is defined similarly to Langford, *et al.* (1998) and Rand (1993). Given homicide cases that exist in both datasets, the SHR data are accurate to the degree that SHR codes agree with CHD codes for the same variable. There are two kinds of disagreement. The category may be missing in the SHR. For example, Langford, *et al.* (1998) and others have pointed out that the SHR has no relationship code for ex-boyfriend or ex-girlfriend. This was one reason for SHR cases not being categorized as intimate partner homicide in the Massachusetts study. With

no code for ex-boyfriend, the reporting officer must chose some other option, such as friend or acquaintance. Similarly, the SHR does not include Latino/Hispanic ethnicity. Another kind of disagreement can be more reasonably called inaccuracy. This occurs when the SHR code exists, but the case is coded differently in the SHR than in the CHD. In the Chicago analysis, we allowed for some leeway in determining whether or not an SHR code was the same as a CHD code. With the victim's age, for example, we separated differences of one or two years from larger discrepancies. Our goal was to focus on degrees of inaccuracy that would be likely to affect policy decisions.

### Case Matching at the Victim Level

The project's first task was to find a unique match in the SHR for each of the 1,784 homicide victims in the CHD booked in 1993 or 1994. Because the CHD contains only criminal homicides, not justifiable homicides nor involuntary manslaughter cases, we did not try to match CHD homicide cases to the 50 SHR justifiable homicides or the 13 SHR involuntary manslaughters for those years. Deleting these 63 cases, there were 1,752 SHR criminal homicides for 1993 and 1994.

On the surface, it would appear to be a simple matter to compare, case by case, each of the Chicago Homicide Dataset cases to an SHR case. However, this undertaking posed formidable methodological challenges -- database organization, case matching and comparative criteria and analysis. Case matching of the victim records in the two datasets was not straightforward, because there is no identification number common to the two. Information that might have been used in lieu of a common identification number (for example, the date and time the lethal incident occurred) does not exist in the SHR. It would not have been logical to have used a combination of victim and offender characteristics, because one goal of the project was to describe the differences between the two datasets in just those characteristics. If we had matched the datasets based on offender characteristics, then they would, of course, match each other in those characteristics.

Staff of the FBI/UCR provided the solution to this dilemma.<sup>1</sup> The paper SHR report forms completed by CPD and sent to the UCR still exist for these years. UCR staff sent the project copies of these paper forms, which we call the "Transfer" files. We attempted to match each of the 1,784 CPD victims to a Transfer file victim, and then to an SHR victim. Unlike the SHR, the Transfer file contains information on time and victim's ethnicity. Therefore, fields used to match a CHD case and a Transfer file case, were date (year, month and day of month), time, victim's age, victim's race/ethnicity and victim's gender. Fields used to match a Transfer case and a SHR case were month, year, victim's age, race and victim's gender, offender's age, race and gender, and situation (multiple or single victims and offenders). In 200 cases, the single "time" recorded in the Transfer file differed from the "death time" in the CHD but was the same as the "injury time." In 121 cases, the victim's age was different in the SHR and the CHD. However, as both Rand (1993) and Langford, *et al.* (1998) found, many of these age differences were only one year.

In matching cases between the Transfer file and the CHD, the goal was to find a unique pair, the one and only one CHD case that matched a given Transfer file case on five key fields. Of the 1,759 Transfer file cases, 667 were the same as one and

---

<sup>1</sup>We are grateful to Vicki Major and Ken Candell for helping us to obtain and interpret these files.

only one CHD case on all five fields: victim's race/ethnicity, victim's gender, victim's age, date and time. To match the remaining 1,092 Transfer cases to one and only one CHD case, we did not always require an exact match on all five fields. Like Langford, *et al.* (1998), we conducted the matching process in stages. As the matched cases were removed, the number of remaining candidates for a unique match declined. Using this method, we found that 164 of the remaining 1,092 cases were the same as only one CHD case on victim's race/ethnicity, victim's gender, victim's age and date, leaving 928 unmatched cases. Of these, 115 cases were the same as only one CHD case on victim's race/ethnicity, victim's gender, date and time, leaving 813 unmatched cases. Of these, 794 cases were the same as only one CHD case on victim's race/ethnicity, victim's gender and date, leaving 19 unmatched cases. Of these, 15 cases were the same as only one CHD case on victim's age, victim's race/ethnicity, victim's gender and month, leaving only four unmatched cases. Two of the four were a complete match with two CHD cases, but not unique. This happened because the two victims had the same demographics and were killed in the same incident, and they were, therefore, the correct match. The final two Transfer file cases were not present in the CHD. In one of these, CPD had initially recorded a body found in the trunk of a car as a Chicago homicide, but further investigation determined that the murder had been committed elsewhere. Thus, the cases should not have been in the SHR as a Chicago case. We were unable to ascertain why the other case was missing in the CHD. It did not have characteristics that matched any CHD case booked in 1993 or 1994, or occurring in 1992, 1993 or 1994.<sup>2</sup>

In summary, we attempted to determine, case by case, whether each of the 1,784 CHD victim homicide records booked in 1993 or 1994 was present among the 1,759 Transfer file cases for those years, and whether each Transfer file case was present among the 1,752 SHR cases for those years. We determined that 1,757 cases were present in the CHD and the Transfer file. Two cases were in the Transfer file but not the CHD, and 27 cases were in the CHD but not the Transfer file. There was no SHR case that did not match a Transfer file case, but there were seven Transfer file cases that were not present in the SHR. Thus, there were 34 CHD cases that were not present in the SHR. Of these, 27 were not present in the Transfer file, and seven more were in the Transfer file but were not present in the SHR.

#### Case Matching at the Offender Level

Because some cases had multiple victims, the 1,750 CHD victim records present in the SHR accounted for 1,674 separate incidents. In 205 of these incidents (12.2%), no offender was positively identified by the police.<sup>3</sup> These 1,469 incidents had as many as eleven offenders per incident, a total of 1,631 offenders known to the police. To examine the completeness and accuracy of offender records, we compiled an offender-level dataset, which contained detailed information from the CHD Offender-

---

<sup>2</sup>The limited information available in the SHR did not allow us to search for a unique match in other years; there could have been thousands of possibilities with the same victim demographics.

<sup>3</sup>Our criterion for "positively identified" is that the case is cleared and police records show the offender's demographics (age, gender and race). "Identified" does not imply "arrested." Some offenders are identified but not arrested. This happens, for example, when an offender commits suicide or dies of other causes.

Level File on each offender and on the related incident and victim(s), plus a dichotomous variable indicating whether or not each of the 1,631 offenders was present or missing in the SHR.

Creating an SHR-missing versus SHR-present variable for each offender was a problem for a few cases, because there is no unique offender-level identifier in the SHR. There was no problem with the 842 incidents with only one offender, or the 775 multiple-offender incidents in which information was either present or missing in the SHR for all of the offenders. However, for the 89 offenders in the 24 multiple-offender incidents in which information was present in the SHR for one or more offender but missing for others, it was necessary make decisions based on limited data. We did this arbitrarily, by designating the first-listed offenders in the CHD file as present in SHR. For example, in a five-offender incident in which the SHR had information for four offenders, we called the first four offenders present and the fifth missing. All five were male and all five were Latino. Ages were close but not identical, ranging from 17 to 20 in the following order: 18, 20, 17, 18, 17. Thus, the final offender, who was age 17, was designated as “SHR missing.” This case is typical of all 24. All multiple offenders in each of these 24 incidents were the same gender and racial/ethnic group as each other. Ages varied slightly, but the order of listing in the CHD is not related to age.

The offender-level dataset developed for this analysis contains information on the 1,631 offenders identified by the police in the Chicago Homicide Dataset, in the 1,469 incidents booked in 1993 or 1994 in which at least one offender was positively identified by the police, and which were matched to an SHR incident. Of the 1,631 offenders, 1,473 are present in the SHR, and 158 (10%) are SHR-missing.

## **Results: Completeness**

### Victim Case Completeness

Of the 1,784 CHD victim cases booked in 1993 or 1994, 1,750 (98%) were present in the SHR. We examined these 34 SHR-missing cases, to determine if they differed in any way from the 1,750 SHR-present cases. The victim’s characteristics, the victim-offender relationship and the circumstances of these cases were the same as the 1,750 SHR-present cases.

Only one characteristic distinguished them – the time lag between injury and death (Table 1).<sup>4</sup> Of the 18 CHD homicide victims who died in the calendar year subsequent to the fatal incident, 15 (83%) were SHR missing, compared to only 18 of the 1,536 victims who died in the same month and calendar year (1%) and one of the 230 victims who died in the same calendar year but the next month (less than 1%). Of the 34 SHR missing cases, eight died in the following calendar year, three died two years later, two died four years later, and two died eight years later. In addition to these fifteen, the bodies of two of the SHR missing victims were found in the year following the death. All of these 17 cases were missing in the Transfer file. One of the missing victims was killed in an incident on December 30 in which another victim was

---

<sup>4</sup>The date of the incident leading to the death and the date of the actual death are both available variables in the CHD, though not in the SHR.

also killed. That victim died immediately, and was recorded in the SHR, but the victim who lived until January 3 was not recorded in the SHR.

**Table 1**  
**Delayed Death and Percent Missing in SHR**

<b>SHR Status</b>	<b>Was Death Delayed After the Fatal Incident?</b>		
	<b>Injury and Death in Same Month and Year</b> (N = 1,536)	<b>Death in Subsequent Month of the Same Year</b> (N= 230)	<b>Death Occurred in a Subsequent Calendar Year</b> (N = 18)
<b>SHR Missing</b>	1.2%	.4%	83.3%
<b>SHR Present</b>	98.8	99.6	16.7
<b>Total</b>	100.0%	100.0%	100.0%

In general, if the victim died in the calendar year following the lethal incident, the case was much more likely to be missing in the SHR. The absolute length of time from injury to death was not important, only whether or not the lag crossed a calendar year. For example, if someone was injured on January 1 and died on December 31 of the same year, 364 days later, that case would be as likely to be present in the SHR as any other case. However, if a victim was injured on December 31 of one year and died on January 1 of the following year, only one day later, that case would be much less likely to be present in the SHR.

Many of the 17 SHR missing victim cases in which there was no lag between injury and death and no delay in finding the body appear to have presented a particularly difficult problem for the police investigators or for the Medical Examiner's Office. Two of the 17 were decomposed or dismembered bodies. Twelve cases appear at the bottom of the paper Transfer file, instead of in the usual sequential order. Finally, seven victims are missing in the SHR but present in the Transfer file. Each of these cases, for various reasons, appears to be due to a data entry error (Patterson, 1997: 22).

#### Incident Case Completeness

Of the 1,750 victim records that exist in both the CHD and the SHR, we do not necessarily know how many separate incidents there were. Accurate information about the number of incidents is important for many policy and research (Block, 1993a). It is also necessary to have an accurate and complete dataset of incidents in order to have an accurate and complete count of offenders. For example, if a group of five teenagers attacked and killed two victims, an accurate offender-level dataset would include each teenager only once, not twice. Population-based rates of homicide offending would be inaccurate otherwise (Block, 1993b).

In the CHD, it is easy to determine the number of incidents, because each victim record contains identification numbers for the victim and for the incident. In order to accurately determine the number of homicide incidents in SHR data, however, we must not only have a complete number of victim records, but the SHR variable delineating the number of victims and offenders (SITUATION) must also be

accurate. Thus, the determination of incident case completeness involves a look at the accuracy of this SHR variable. To do this, we analyzed victim count and offender count information in SITUATION separately.

According to CHD information, the 1,750 victim records that exist in both the CHD and the SHR account for 1,674 separate incidents. Of these, 1,612 were single-victim incidents, 54 were two-victim incidents, five were three-victim incidents, two were four-victim incidents and one was a seven-victim incident. We found that the SHR records for the 1,750 victims contained accurate information about the number of victims for all of the eight incidents with three or more victims. (SITUATION indicated "multiple victims," and the number of victims in the file was the same as the number in the CHD.)

However, there were nine victims in the SHR who were classified as single-victim, but were in a two-victim incident according to the CHD information. These nine include the second victim of a two-victim incident that had occurred in 1990, and eight victims in four two-victim incidents in the 1993-1994 SHR dataset. In each of the four incidents, one of the victims died at the scene or soon afterward, while the second lingered for three days to 28 days before dying. Therefore, the Transfer file records for the two victims were physically separated on different pages. Without a unique incident identifier, it would have been almost impossible for the SHR data entry person to have cross-referenced the two victims. Since there appeared to be no record for a second victim, the incident was entered as a single-victim incident.

In addition, in six CHD incidents, the SHR contained errors in the number of offenders. In five cases, there was only one offender according to CHD information, but the SHR contained information for two offenders. Four of these were two-victim, one-offender incidents. It is possible that the SHR data entry operator mistakenly entered 1 in the "Additional Offender" field, which is next to the "Additional Victim" field. The fifth incident was a one-victim, one-offender incident in the CHD. Demographic information about the single offender was duplicated in SHR fields for a second offender. The sixth incident was an unknown-offender incident in the CHD, with all offender information missing. In the SHR, offender demographic information for this case were duplicated from the case immediately preceding it in the Transfer file.

Thus, there are some errors in the SHR counts for the 1,750 homicide incidents in the CHD. Because the SHR counted nine victims as having been murdered in single-victim incident instead of a double-victim incident, the number of incidents in these cases was doubled. In six incident, the number of offenders in the SHR file was doubled. Although these errors seem small in comparison to the total number of incidents, they could have been prevented if there had been unique identification numbers for the victim, the incident and the offender.

### Offender Case Completeness

Of the 1,631 offenders known to the police in the 1,469 CHD incidents that exist in the SHR and in which police have identified at least one offender, the SHR contains information about 1,473 (90%). Information about 158 (9.7%) offenders is missing (table 2) . Of the 1,469 CHD incidents, 93 (6%) had no offender information in the SHR, including 72 of the 842 single-offender incidents (9%), 16 of the 186 two-offender incidents (9%), two of the 68 three-offender incidents (3%), two of the 26 four-offender incidents (8%), and one of the ten five-offender incidents. For an

additional 24 CHD incidents (1.4%), the SHR contained information about one or more of the offenders, but was missing information about others. This includes five of the 186 two-offender incidents (3%), 11 of the 68 three-offender incidents (16%), five of the 26 four-offender incidents (19%), two of the ten five-offender incidents, and one of the four six-offender incidents.

Table 2  
**Offenders Missing in the SHR, Compared to the CHD**

Number of Identified Offenders in Incident	Total Offenders	Status in the SHR		
		Number Present	Number Missing	Percent Missing
1	842	770	72	8.6%
2	374	332	42	11.2
3	204	185	19	9.3
4	132	115	17	12.9
5	30	23	7	23.3
6	24	23	1	4.2
7	14	14	0	.0
8	0	0	0	na
9	0	0	0	na
10	0	0	0	na
11	11	11	0	.0
<b>Total</b>	<b>1,631</b>	<b>1,473</b>	<b>158</b>	<b>9.7%</b>

Overall, even in the CHD homicide incidents in which victim and incident information is present in the SHR, 7% percent were missing information about at least one of the offenders known to the police. Of the CHD homicide offenders known to the police, SHR records are missing for almost ten percent, even when the incident itself is present in the SHR. This is a much higher level of incompleteness than the 2% SHR-missing victim records.



## Results: Accuracy

As an indicator of accuracy, we conducted a detailed analysis of variables in each of the 65 CHD victim-level cases booked in January, 1993, and present in the SHR. In this month, 28 of the CHD cases contained at least one variable in which codes differed in the CHD and the SHR (Patterson, 1997a: 14).

Age. In 15 of the cases, the victim or offender age was different in the SHR compared to the CHD. Nine of these were a difference of one year. All of these differences occurred in the Transfer file.

Race/Ethnicity. The racial/ethnic group of victim and offender in the CHD delineates non-Latino white, non-Latino black, Latino white and Latino black. This information is retained in the Transfer file. However, in the SHR in 1993 and 1994, the Ethnic Origin variable was coded "unknown" for all cases. Under the Race variable, victims or offenders who were categorized as Latino white in the Transfer file were coded "white". In these 65 cases, there were 27 in which a Latino white victim or offender was coded as race white, ethnic origin unknown in the SHR.

Gender. All 65 cases had accurate information on victim and offender gender.

Situation. In nine SHR cases, the number of offenders was coded as unknown, when the CHD contained that information.

Weapon. The CHD has more detailed weapon codes than the SHR. For example, it differentiates between semi-automatic and non-automatic handguns. There were 19 CHD cases of the 65 in which the SHR code was correct, but less specific than the CHD code. These included 18 cases in which the CHD code for semi-automatic handgun was recorded in the SHR as handgun, and one case in which the CHD code for "unknown firearm" was recorded in the SHR as "unknown weapon." In addition, there were four cases in which Weapon was coded inaccurately in the SHR. In three of these, "unknown firearm" was recorded in the SHR as "handgun," and in one case, "club" was recorded in the SHR as "fists: personal weapon, hands, feet, etc.".

Victim and Offender Relationship. In 22 of the 65 cases, there were differences between the CHD and the SHR relationship code. Many of these were due to the more specific codes available in the CHD. CHD codes of proprietor, customer, and drug pusher were recorded as stranger in the SHR, and CHD codes of client, employer and ex-boyfriend were recorded as acquaintance in the SHR. One CHD case of a stepfather, one case of an acquaintance, and one case of a stranger were recorded as unknown in the SHR. Of two CHD cases of a homosexual domestic relationship, one was recorded as boyfriend and one as unknown in the SHR. Of eight CHD cases of a rival gang member relationship, the SHR recorded three as stranger and five as unknown. A CHD case of a gang member killing a non-gang victim was recorded as stranger in the SHR.

Circumstances. In 23 of the 65 cases, there were differences between the CHD and the SHR circumstance code. Three cases of street gang were recorded as narcotics, gangland killing and unknown. Of nine cases of domestic argument, one was recorded as an argument due to alcohol, one as an argument over money, and seven as "other argument." Four CHD cases coded traffic argument, interference in felony, retaliation, and argument over attempted theft were recorded as "other argument" in the SHR. Two CHD cases of "argument over theft" were recorded as "argument over money" in the SHR.

Though there were a number of inaccuracies in the 65 cases booked in January, 1993, most of these inaccuracies were due to SHR codes that are not as specific as CHD codes, or to case information that had been updated in the CHD but not in the SHR. In general, the inaccuracies were not due to systematic data entry errors. Many of the inaccuracies occurred in just a few categories that do not exist in the SHR, including street gang relationship (rival gang members and a non-gang member killed by a gang member), street gang-related motive, Latino ethnicity, semi-automatic firearm, homosexual domestic relationship, and ex-boyfriend.

### **Results: Bias in SHR Offender Records**

Of the 1,631 offenders present in the Chicago Homicide Dataset, 158 are not present in the SHR. With this dataset, we can now explore the questions initially posed by BJS:

1. Do the 158 offenders who are missing in the SHR differ significantly from offenders who are not missing, and if so, how are they different?
2. Do offender characteristics in the complete offender dataset (the 1,631 CHD cases) differ from offender characteristics in the SHR dataset (1,473 cases)?

#### Differences between Offenders Present in SHR and Offenders Missing in SHR

Of the 1,631 Chicago homicide offenders known to the police in 1993 and 1994, who is more likely to be missing in the SHR? We found significant differences between the two groups in the racial/ethnic group of the offender, whether the offender had a prior record, and the offender's age. We also found that offenders who used a firearm, or who committed a street gang-related homicide, were more likely to be SHR-missing.

Prior Record. Whether or not the offender had a prior record is a strong indicator that the offender will be missing in the SHR. Of the 1,252 offenders with an arrest record, 10.9% were SHR-missing, compared to 5.5% of the 379 offenders with no record ( $t = -3.758$ ;  $p < .0005$ ). The correlation between having a prior record and being SHR-missing was .077 ( $p < .01$ ). There was no significant difference, however, between those with violent prior records (10.7% missing in SHR) and those with non-violent records (13.0% missing).

Race/Ethnicity and Gender. Compared to other racial/ethnic and gender groups, Latino male offenders are more likely to be missing and black female offenders are more likely to be present in the SHR. Of the 286 Latino males, 14.3% were SHR-missing, compared to 8.7% of other offenders ( $t = -2.566$ ;  $p = .011$ ). Only one of the 112 black women offenders was missing in the SHR ( $t = 7.384$ ;  $p < .0005$ ).<sup>5</sup> Correlations between membership in a racial/gender group and SHR-missing are significant only for Latino males and black females (Table 3). Black and non-Latino white males are no more or less likely to be SHR-missing than others.

Age. The mean age of SHR-present offenders (23.6 years) is slightly but significantly ( $t = 4.017$ ;  $p < .0005$ ) older than the mean age of SHR-missing offenders

---

<sup>5</sup>There were only ten Latina females and four non-Latina white females in the dataset. One of the Latina offenders was SHR-missing, and none of the four white female offenders.

(21.5 years). However, correlations between membership in a specific age group and SHR-missing status are significant only for a few age groups (table 4). While 10.7% of the 665 offenders aged 15 to 19 and 12.9% of the 403 offenders aged 20 to 24 are SHR-missing, none of the 70 offenders aged 45 to 74 and only two (4.7%) of the 43 offenders aged 40 to 44 were SHR-missing. However, the correlation between membership in an age group and being SHR-missing is significant only for the 20-to-24-year old group ( $r=.062$ ;  $p=.012$ ) and those aged 50 to 74;  $p= -.057$ ;  $p=.021$ ).

Table 3  
Correlations: SHR Missing and Racial/Gender Group

Race/Ethnicity and Gender	Pearson Correlation with SHR Missing <sup>6</sup>	Sig. (2-tailed)
Black male	.007	NS
Black female	-.081**	.001
Latino male	.072**	.003
Latina female	.001	NS
White male	-.035	NS
White female	-.016	NS

Table 4  
Correlations: SHR Missing and Age Group

Age Group	Pearson Correlation with SHR Missing <sup>7</sup>	Sig. (2-tailed)
Age 15 to 19	.028	NS
Age 20 to 24	.062*	.012
Age 25 to 34	-.026	NS
Age 50 to 74	-.057*	.021

Motive and Circumstance. Offenders in cases with certain motivations or situations were significantly more or less likely to be SHR-missing. The dataset contains only eight offenders involved in a contract killing, but five of the eight were SHR-missing. Of the 178 offenders in an armed robbery homicide, 22 (12.4%) were SHR-missing. Offenders in a street gang-motivated homicide were more likely to be SHR-missing than others, as were offenders in an instrumental homicide.<sup>8</sup> Of the 606 offenders involved in a street gang homicide, 12.0% were missing in SHR (Chi square

<sup>6</sup>For a correlation matrix of these variables, see Appendix I.

<sup>7</sup>For a correlation matrix of these variables, see Appendix I.

<sup>8</sup>In an instrumental homicide, the offender's primary and immediate motive is to get money or property.

$p=.015$ ;  $\text{Gamma}=.205$ ,  $p=.017$ ). Being involved in an instrumental homicide was not a significant predictor of whether the offender would be SHR-missing.

Only three of the 122 offenders involved in an intimate partner homicide were missing in the SHR (Chi square  $p = .002$ ;  $\text{Gamma} = -.639$ ,  $p < .0005$ ). Similarly, none of the 27 offenders involved in a child abuse homicide was SHR-missing.

Substance Use. The Chicago Homicide Dataset differentiates between drug motive and drug use in a homicide. A homicide is motivated by drugs when the homicide was committed to get drugs or to get money for drugs, when the homicide grew out of an argument about drugs, when the homicide was committed in order to further a drug business, or when drugs motivated the homicide in some other way. We found that whether or not the offender was involved in a drug-motivated homicide made no difference in whether the offender would be missing in the SHR. This was true regardless of the type of motivation or whether the evidence of drug motivation was positive or circumstantial.

Although drug motivation makes no difference, we found that drug or liquor use increases the chance that an offender will be present, not missing, in the SHR. Only one of the 69 offenders involved in an incident in which someone was using drugs (the offender, the victim, or both) was missing in the SHR (Chi square  $p = .018$ ;  $\text{Gamma} = -.767$ ,  $p < .0005$ ). Similarly, of the 412 offenders in an incident in which someone was using liquor, only 5.6% were missing (Chi square  $p = .001$ ;  $\text{Gamma} = -.356$ ,  $p < .0005$ ).

Weapon. The 1,191 offenders who committed the homicide with a firearm were much more likely to be SHR-missing than the 440 offenders who did not (11.8% versus 4.1%), and the difference was highly significant (Chi square  $p < .0005$ ;  $\text{Gamma} = .515$ ,  $p < .0005$ ;  $t = 4.673$ ,  $p < .0005$ ). Of the firearms, offenders using a semi-automatic weapon (12.1%), another handgun (12.0%), or a shotgun (19.0%) were more likely to be SHR-missing than the 20 offenders using a non-automatic rifle (5.0%). The 194 offenders who used a knife or other sharp instrument were much less likely to be SHR-missing (2.6%).

Number of Victims and Offenders. For offenders who are SHR-missing, the mean number of victims in the incident was exactly the same as for offenders who are SHR-present (1.05), and the mean number of offenders in the incident is almost the same (2.04 for SHR-missing offenders and 1.98 for SHR-present offenders). The correlations are not significant and only .011 (number of offenders) and -.002 (number of victims).

Of the 70 offenders who killed multiple victims, 11.4% were SHR-missing, compared to 9.6% of offenders who killed only one victim. Of the 789 offenders in multiple-offender homicides, 10.9% were missing, compared to 8.6% of the offenders in single-offender homicides. Neither Chi square nor Gamma is significant in either case.

Multivariate Analysis. Thus, a number of offender and incident characteristics have significant, though often small, relationships to whether or not the offender will be missing in the SHR. Factors *increasing* the risk that the offender will be SHR-missing are being a Latino male, being age 20 to 24, having a prior arrest record, being involved in a street gang-motivated homicide, and using a firearm other than a rifle. Factors significantly *decreasing* the risk that the offender will be SHR-missing are being a black female, being age 40 or older, being involved in an intimate partner

homicide, being involved in a homicide in which either the victim or offender was using drugs or liquor, and using a knife or sharp instrument.

However, many of these factors are highly related to each other. For example, the offender's age is correlated with whether or not the offense was street gang-related ( $r = -.391$ ). To determine how all of these aspects of the offender and the homicide situation might work together to affect whether or not the offender will be SHR-missing, we conducted a multivariate stepwise regression analysis. This resulted in a model with four variables, firearm use, prior record, Latino male and liquor use (table 5). After these four were entered in a stepwise regression, none of the other variables was significant. This model explains little of the variance in SHR-missing. The  $r$  is .163 and the  $R^2$  is only .027.

Table 5  
Multiple Regression, Explaining SHR-Missing Status

Variable in Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t value	Sig.
Constant	1.009	.020		49.819	.000
Firearm?	.06322	.017	.095	3.780	.000
Prior record?	.05650	.017	.081	3.293	.001
Latino male?	.04882	.019	.063	2.544	.011
Liquor use?	-.04240	.017	-.062	-2.499	.013

Logistic regression, which is more appropriate because SHR-missing is a dichotomy, produces similar results, except there are five variables in the model. Whether or not the offender is a Black female is added to the other four (table 6).

Table 6  
Logistic Regression Model

Variable	B	S.E.	Wald	df	Sign.	R	Exp (B)
Constant	-3.4601	.3355	106.3665	1	.0000		
Prior record?	.7339	.2452	8.9582	1	.0028	.0819	2.0832
Firearm?	.8620	.2630	10.7428	1	.0010	.0918	2.3678
Latino male?	.4441	.1982	5.0220	1	.0250	.0540	1.5591
Black female?	-1.8223	1.0001	3.3199	1	.0684	-.0357	.1617
Liquor use?	-.5827	.2374	6.0254	1	.0141	-.0623	.5584

These models delineate a group of offenders who are significantly much more likely to be missing in the SHR. Of the 143 offenders who meet all four of the criteria in the model described in table 5 (Latino males with a prior record who used a firearm in an incident where liquor was used), 18.9% were SHR-missing, compared with 8.8% of the other 1,488 offenders (Chi square  $p < .0005$ ; Gamma = .414,  $p = .003$ ).

Bias in SHR Offender Data

Even though certain offender and incident characteristics, taken singly and together, make it much more likely that a particular offender will be missing or not missing in the SHR, the aggregate total of all of these “SHR inclusion decisions” does not have a large effect on the overall characteristics of the SHR dataset (Table 7).

Table 7  
**Differences between the Chicago Homicide Dataset and the SHR:  
 Offender Characteristics**

Offender Characteristic	CHD Dataset (n = 1,631)		SHR Dataset (n = 1,473)	
	N	%	N	%
<b>Has a Prior Record</b>	1,252	76.8%	1,115	75.7%
<b>Black male</b>	1,641	69.2%	1,018	69.1%
<b>Black female</b>	112	7.5%	111	6.9%
<b>Latino male</b>	286	17.5%	245	16.6%
<b>Latino female</b>	10	.6%	9	.6%
<b>White male</b>	79	4.8%	75	5.1%
<b>White female</b>	4	.2%	4	.3%
<b>Aged 10 to 14</b>	70	4.3%	68	4.6%
<b>Aged 15 to 19</b>	665	40.8%	594	40.3%
<b>Aged 20 to 24</b>	403	24.7%	351	23.8%
<b>Aged 25 to 34</b>	308	18.9%	283	19.2%
<b>Aged 35 to 49</b>	137	8.4%	129	8.8%
<b>Aged 50 to 74</b>	48	2.9%	48	3.3%
<b>Stranger to victim<sup>9</sup></b>	177	10.9%	163	11.1%
<b>Drug use, victim or offender</b>	69	4.2%	68	4.6%
<b>Liquor use, victim or offender</b>	412	25.3%	389	26.4%

For example, having a prior arrest record is very important in determining whether an offender will be missing in the SHR. Offenders with a record are over twice as likely to be SHR-missing as offenders without a record. Almost 11% of offenders with a record were SHR-missing, compared to 5.5% of other offenders. Despite this strong relationship, however, there is a very small difference between the

<sup>9</sup>All the offenders in the incident were strangers to the victim.

76.8% of offenders in the CHD who have a prior arrest record versus the 75.7% percent of offenders in the SHR who have a prior arrest record (see Table 7).

Similarly, even though only three of the 122 offenders involved in an intimate partner homicide was SHR-missing, there is very little difference in the proportion of intimate partner homicides in the two datasets (Table 8). While 7.5% of the 1,631 homicide offenders in the CHD were involved in an intimate partner homicide, 8.1% of the 1,473 offenders in the SHR were involved in an intimate partner homicide.

Table 8  
**Differences between the Chicago Homicide Dataset and the SHR:  
Incident Characteristics**

Incident Characteristic	CHD Dataset (n = 1,631)		SHR Dataset (n = 1,473)	
	N	%	N	%
<b>Drug-related motive<sup>10</sup></b>	233	14.3%	215	14.6%
<b>Street gang-related motive</b>	606	37.2%	533	36.2%
<b>Instrumental<sup>11</sup></b>	308	18.9%	275	18.7%
<b>Child abuse</b>	28	1.7%	28	1.9%
<b>Murder of an intimate partner<sup>12</sup></b>	122	7.5%	119	8.1%
<b>Sexual assault</b>	15	.9%	10	.7%
<b>Multiple offenders</b>	789	48.4%	703	47.7%
<b>Multiple victims</b>	70	4.3%	62	4.2%
<b>Occurred at home</b>	356	21.8%	341	23.2%
<b>Occurred at a tavern, liquor store</b>	26	1.6%	22	1.5%
<b>Occurred on the street</b>	706	43.3%	631	42.8%
<b>Firearm, total</b>	1,091	73.0%	1,051	71.4%
<b>High caliber semi-automatic weapon</b>	264	16.2%	230	15.6%
<b>Handgun, not semi-automatic</b>	392	24.0%	345	23.4%
<b>Knife or sharp instrument</b>	194	11.9%	189	12.8%
<b>Club or blunt instrument</b>	112	6.9%	109	7.4%
<b>Hands, fists, feet</b>	85	5.2%	78	5.3%

<sup>10</sup>Includes selling, an argument over possession, getting money for drugs, or other drug involvement, but not drug use only. Includes incidents with circumstantial evidence of drug motivation.

<sup>11</sup>Definition of instrumental: the offender's primary and immediate motive is to obtain money or property.

<sup>12</sup>This includes multiple-offender cases in which the victim (or one of the victims) was the intimate partner of one of the offenders.

Therefore, research conclusions or policy decisions about offenders would not necessarily be erroneous if they were based on the SHR offender data, provided that the conclusion and decisions were limited to the kinds of gross totals shown in Tables 7 and 8. A difference, for example, between using an offender-level dataset in which 16.6% of the offenders are Latino men (the SHR), and using a dataset in which 17.5% of the offenders are Latino men, would probably not be a big enough difference to affect most policy decisions.

On the other hand, some research conclusions or policy decisions might depend upon detailed analysis, where a single case or a small number of cases would make a difference. In such a situation, relying on the SHR data could lead to erroneous conclusions. This would be especially true for combinations of characteristics that are particularly prone to be missing or to be present in the SHR. For example, an SHR analysis of weapon use by racial/ethnic group, comparing gang-related to other types of homicide, would be suspect.

### **Conclusions, Causes, and Remedies**

If the SHR contains systematic bias in offender information, how can the completeness of SHR offender information be improved? To answer this question, we compared characteristics of the cases and the investigation process in offender-missing and offender-present incidents. CHD victims and homicide incidents that were missing completely from the SHR differed from those that were not missing in that almost all of the missing cases had a time lag between injury and death that spanned a calendar year. This was not true, however, of the homicides in which the incident was present in the SHR, but the offender information was missing.

There was a significant difference, however, for the calendar-year lag between the date of injury and the date of arrest or other clearance. Of the 146 offenders in incidents cleared in the calendar year following the incident, 71.2% are missing in the SHR. Of the six offenders in incidents cleared two calendar years following the year of the incident, all are SHR-missing. In contrast, only 2.9% of the 1,467 offenders in incidents cleared in the same calendar year were missing in the SHR.

This suggests obvious possibilities for improving the quality of SHR data. To the degree that it is possible to improve the extent to which SHR data can be updated to include offender information for incidents cleared after the data are reported to the SHR, and for multiple-offender incidents in which an additional offender is arrested or otherwise identified, the bias in SHR offender data will be improved.



## References

- Bailey, William C. and Ruth D. Peterson (1995). Gender inequality and violence against women: The case of murder. Pp. 174-205 in John Hagan and Ruth D. Peterson (Eds.), *Crime and Inequality*. Stanford, CA: Stanford University Press.
- Block, Carolyn Rebecca (1993a). Organizing a dataset to support analysis of multiple units: victim, incident and offender risk. Pp. 209-212 in Carolyn Rebecca Block and Richard L. Block (Eds.), *Questions and Answers in Lethal and Non-Lethal Violence 1993*. NCJ 147480. Washington, D.C.: National Institute of Justice.
- Block, Carolyn Rebecca (1993b). Lethal violence in the Chicago Latino Community. Pp. 267-343 in Anna V. Wilson (ed.), *Homicide: The Victim/Offender Connection*. Cincinnati: Anderson Publishing Co.
- Cantor, David and Lawrence E. Cohen (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.
- Chu, Rebekah, Craig Rivera and Colin Loftin (1998). Herding and homicide: An examination of the Nisbett-Reaves hypothesis. Violence Research Group Discussion Paper 17, University of New York at Albany.
- Fox, James Alan (1996). *Uniform Crime Reports [United States]: Supplementary Homicide Reports, 1976-1994. ICPSR version*. Boston, MA: Northeastern University, College of Criminal Justice [producer]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].
- Langford, Linda, Nancy Isaac and Stacey Kabat (1998). Homicides related to intimate partner violence in Massachusetts: Examining case ascertainment and validity of the SHR. *Homicide Studies* 2(4): 353-377.
- Lott, John R., Jr., and David B. Mustard (1997). Crime, deterrence and right-to-carry concealed handguns. *Journal of Legal Studies* 26 (1):1-68.
- Maltz, Michael D. (1999). Bridging Gaps in Police Crime Data. BJS Executive Summary, NCJ 177615. Washington, D.C.: Bureau of Justice Statistics.
- Maltz, Michael D. (1999). Bridging Gaps: Estimating Crime Rates from Incomplete Police Data. Manuscript. Washington, D.C.: Bureau of Justice Statistics.
- Maxfield, Michael G. (1989). Circumstances in Supplementary Homicide Reports: Variety and validity. *Criminology* 27(4) :671-694.
- Miller, Louise S. (1983). *Murder in Illinois: 1973 to 1982*. Chicago, IL: Illinois Criminal Justice Information Authority.
- Miller, Louise S. and Carolyn Rebecca Block (1983). *Illinois Murder Victim Data, 1973 to 1981: Guide to Quality, Availability and Interpretation*. Chicago, IL: Illinois Criminal Justice Information Authority.
- Miller, Louise S. and Carolyn Rebecca Block (1982). *Victim-Level Murder File: Users Guide and Codebook*. Chicago, IL: Illinois Criminal Justice Information Authority.
- Patterson, Thomas D. (1997). SHR Data Quality Project Deliverable 4: Merged Database and Summary Findings. Report to BJS. Manuscript.
- Poggio, Eugene C., Stephen D. Kennedy, Jan. M. Chaiken and Kenneth E. Carlson (1985). *Blueprint for the Future of the Uniform Crime Reporting Program*.

- Washington, D.C.: Bureau of Justice Statistics and Federal Bureau of Investigation.
- Rand, Michael R. (1993). The study of homicide caseflow: Creating a comprehensive homicide dataset. Pp. 103-118 in Carolyn Rebecca Block and Richard L. Block (Eds.), *Questions and Answers in Lethal and Non-Lethal Violence 1993*. NCJ 147480. Washington, D.C.: National Institute of Justice.
- Riedel, Marc (1999). Sources of homicide data. In M. Dwayne Smith and Margaret A. Zahn (Eds.), *Homicide Studies: A Sourcebook of Social Research*. Thousand Oaks, CA: Sage.
- Riedel, Marc (1998). Counting stranger homicides. *Homicide Studies* 2(2): 206-219.
- Riedel, Marc (1993). *Stranger Violence: A Theoretical Inquiry*. New York: Garland.
- Riedel, Marc (1990). Nationwide homicide data sets: An evaluation of the Uniform Crime Reports and the National Center for Health Statistics data. Pp. 175-205 in Doris L. MacKenzie Layton, Phyllis Jo Baunach and Roy R. Roberg (Eds.), *Measuring Crime: Large-Scale, Long-Range Efforts*. Albany: State University of New York Press.
- Rokaw, William M., James Mercy and Jack Smith (1990). Comparability and utility of national homicide data from death certificates and police records. *Public Health Reports* 105: 447-455.
- Snyder, Howard (1996). *National Juvenile Violent Crime Trends, 1990-1994*. Pittsburgh: National Center for Juvenile Justice.
- Williams, Kirk and Robert L. Flewelling (1987). Family, acquaintance, and stranger homicide: Alternate procedures for rate calculations. *Criminology* 25: 543-560.
- Zahn, Margaret A. and Marc Riedel (1983). National versus local data sources in the study of homicide: Do they agree? Pp. 103-120 in G. P. Waldo (Ed.), *Measurement in Criminal Justice*. Beverly Hills: Sage Publications.