Within and Inter-Institutional Differences Between Death Certifiers on Autopsy Conclusions

Ming-Li Hsieh, PhD¹
and Melanie-Angela Neuilly, PhD²

Abstract
This study seeks to establish whether medico-legal practitioners differ in their autopsy conclusions within and across medico-legal institutions. Data include 459 violent deaths (homicides, suicides, and accidents) autopsy reports written by more than 20 death certifiers from four medico-legal institutions in two countries (France and the United States). Multinomial models show that compared with accidental deaths, weapon use and decedents' characteristics both influence a homicide verdict, but not a suicide one. In addition, French practitioners are more likely than Americans to reach a conclusion of homicide or suicide compared with accident, and homicides are more likely to be certified by male practitioners.

Keywords
mortality statistics, violent deaths, death certifiers, international comparison

Introduction
Mortality Statistics as Biased
The debate surrounding the reliability and validity of mortality statistics is not about whether they are indeed valid and reliable, but rather whether their

¹University of Wisconsin-Eau Claire, Eau Claire, WI, USA
²Washington State University, Pullman, WA, USA

Corresponding Author:
Melanie-Angela Neuilly, Assistant Professor, Department of Criminal Justice and Criminology, Washington State University, P.O. Box 644872, Pullman, WA 99164-4872, USA.
Email: m.neuilly@wsu.edu
lack of validity and reliability is so complete that they should not be used at all. The research indeed very consistently identifies evidence of various types of errors in mortality statistics, particularly when it comes to suicides and deaths of undetermined intent (Breiding & Weirsema, 2006; Bugeja, Clapperton, Killian, Stephan, & Ozanne-Smith, 2010; Goodin & Hanzlick, 1997; Hanzlick & Goodin, 1997). As a matter of fact, Hanzlick, Hunsaker, and Davis (2002) emphasize the absence of “right or wrong” conclusions in death certification and provide a guide with regard to “degrees of certainty” stressing the flexible nature of the decision-making process at the source of mortality statistics. Based on the findings of their meta-analysis, however, Tollefsen, Hem, and Ekeberg (2012) conclude that more systematic reviews are needed, as the existing body of findings range broadly in estimates of misclassifications (from none or very little up to more than 30% with regard to suicide underreporting).

Pioneers such as Douglas (1967) argue that these limitations are evidence that as socially constructed artifacts, suicide statistics reflect the institutions that create them as well as the social context in which these institutions function rather than the behaviors they aim at measuring. This, in turn, implies that mortality statistics are of no use for theory testing and prevention strategy development (Stengel, 1968; Whitt, 2006). Others, however, are more moderate in their findings, arguing that there are ultimately too few misclassified cases to influence mortality research (O’Carroll, 1989; Pescosolido & Mendelsohn, 1986; Sainsbury & Jenkins, 1982).

Factors Influencing Death Certification

Whether mortality statistics can be used in research hinges on identifying their sources of bias. Generally speaking, the extant research has focused on decedents’ characteristics, incidents characteristics, and institutional characteristics.

Gender, race, and age of decedents all play a part in influencing the medico-legal decision-making process. Indeed, while, from a strictly epidemiological perspective, men die at higher rates and earlier than women, across almost all types of natural as well as external causes of death, and across time and space (Dannenberg, Baker, & Li, 1994; Rajaratnam et al., 2010; Schnitzer & Runyan, 1995; Sorenson, 2011), women’s deaths are more likely to be classified as of undetermined intent than suicides (Linsley, Schapira, & Kelly, 2001; Sorenson, Haikang, & Kraus, 1997). With regard to race, the wide variation in suicide rates between Whites, Blacks, and Hispanics is a starting point when it comes to understanding how race may influence manner of death conclusions (Franzini, Ribble, & Keddie, 2001; Phillips & Ruth, 1993; Rockett, Samora, & Coben, 2006; Rockett et al., 2010; Warshauer & Monk,
Deaths of African Americans, Asians, and Native Americans are more likely to be classified as of undetermined intent (Rockett et al., 2010; Sorenson et al., 1997). Finally, deaths of younger individuals are more likely to be misclassified, particularly when it comes to suicide (Linsley et al., 2001; Platt, Backett, & Kreitman, 1988; Taylor, 1982).

Certain types of deaths are more likely to be seen as undetermined or open verdicts (Lindqvist & Gustafsson, 2002; Linsley et al., 2001; Lunetta, Smith, Penttila, & Sajantila, 2003; Sorenson et al., 1997; Stanistreet, Taylor, Jeffrey, & Gabbay, 2001). Unsurprisingly, the presence of a suicide note or a known suicidal intent increases the likelihood that a death is classified as a suicide (Lindqvist & Gustafsson, 2002; Linsley et al., 2001; Lunetta et al., 2003; Stanistreet et al., 2001), and the quantity and quality of information accessible to medico-legal practitioners also has an impact on medico-legal practitioners’ ability to draw strong conclusions (O’Carroll, 1989; Jobes, Berman, & Josselson, 1987).

The cultural and institutional context in which mortality statistics are created has probably been the largest area of interest in the field. There are indeed many structural differences that may influence the medico-legal decision-making process and therefore mortality statistics. Such differences start with cultural and legal definitions of violent deaths (Bertolote & Fleischmann, 2002; Chishti, Stone, Corcoran, Williamson, & Petridou, 2003; Hendin, 1964, 1969; Neuilly, 2007, 2011, 2013; O’Carroll, 1989). The differences between coronial and medical models of medico-legal institutions have of course been of central interest to the field. Varnik and colleagues (2010) provide a good overview of such differences, specifically with regard to the standards used (legal vs. medical) in the decision-making processes (see also, Jougla et al., 2002; Neuilly, 2011; Rockett, 2010). However, even within one type of system (Studdert & Cordner, 2010; Walters, Bugeja, Spittal, & Studdert, 2012), or one institution, changes in policies and practices can affect statistics (see Whitt, 2006 and his presentation of the New York City medical examiner’s office). Overall, the issue is one of lack of uniformity and multiplicity of actors involved (De Leo et al., 2010).

**Objectives**

At the heart of those mortality statistics problems, and often left under-examined (aside from Goodin & Hanzlick’s, 1997 landmark study), is the decision-making process leading individual medico-legal practitioners, those in charge of certifying deaths, to classify deaths as natural, suicidal, accidental, homicidal, or undetermined. These medico-legal practitioners occupy a peculiar place in Western society. Whether they are medical examiners or coroners,
they have been placed on the pedestal of science and made the arbiters of truth in death, firmly establishing their conclusions as definitive, objective, and certain, one of the sources of legitimacy for their forensic authority (Timmermans, 2006). Research on medico-legal practices and the certification of the manner of death has repeatedly shown, however, that there is no such thing as definitive, objective, and certain conclusions when it comes to establishing manner or cause of death (see previously cited sources, but also Goodin & Hanzlick, 1997; Hanzlick & Goodin, 1997; Hanzlick et al., 2002).

Method

Here, using primary source autopsy report data from four medico-legal offices in two countries, we examine the impact of individual characteristics of medico-legal practitioners and compare them with institutional differences between death certifying sites on manner of death classification.

Data from this study were extracted from a set of 720 death investigation and autopsy report files collected by the second author, during field research at four medico-legal sites. The two American sites include a large regional medical examiner’s office from an urban area on the East Coast and a mid-sized county coroner’s office located in the Intermountain West. The two French sites are both medico-legal institutes attached to university hospitals, one in a mid-sized city with a large university in the Western region and the other being in a larger city with a proportionately smaller university in the Southern region. Site selection followed a combination of convenience and purposive sampling strategies, aiming to provide as much representativeness as possible outside the bounds of probability sampling. Information collected on each death includes decedents’ and incidents’ characteristics, as well as investigation and post mortem characteristics and results.

For purposes of this study, we only selected violent deaths (which amounted to 459 cases), making our outcome measure tri-categorical (homicide, suicide, accident). To ascertain the relative influence of individual medico-legal practitioners’ characteristics, our two primary independent variables were dummy-coded (American vs. French and male vs. female practitioners). Control variables included the mainstay of the extant literature: decedents’ demographics (gender, race, age) and incidents’ characteristics (weapon use, setting, and whether the decedent was pronounced dead at scene), all of which (except age) were dichotomized. We thus used a multinomial logistic regression model to examine the impact of medico-legal practitioners’ characteristics across the three-category outcome variable.
Table 1 presents descriptive statistics for the characteristics of the entire sample (N = 459). Overall, 43% of medico-legal practitioners were American and 56% of them were male. In terms of manner of death, the majority of cases were accidents (45%), followed by suicides (31%) and homicides (22%). Roughly 36% of incidents involved weapons (whether firearm, blunt, or sharp objects). Almost half of the decedents were found dead in residences and two thirds were pronounced dead at the scene. The majority of decedents were male, White, and, on average, 45 years old.

Table 2 presents the results of the multinomial logistic regression, which tests the differences by medico-legal practitioners across a three-category manner of death outcome variable. The findings reveal a good model fit (−2Log Likelihood = 262.926, χ² = 91.635, p < .001, Nagelkerke R² = .445). Overall, our results align with the extant literature. In terms of circumstances, cases classified as homicides are more likely to involve a weapon (odds ratio [OR] = 2.843) and those decedents are less likely to be pronounced dead at the scene (OR = 0.239) compared with cases classified as accidental deaths. It is the reverse for suicides, for which decedents are more likely to be pronounced at the scene compared with accidents (OR = 4.101). Compared with accidental deaths, both older decedents (OR = 0.239) and White decedents (OR = 0.089) were less likely to be victims of homicides. Decedents’ characteristics did not influence whether a death was classified as a suicide versus an accident.

With regard to within and inter-institutional differences between medico-legal practitioners, our analyses uncovered some promising results. First,
there were significant differences between the American and French medico-legal practitioners pointing to inter-institutional differences. Indeed, compared with the cases certified as accidental deaths, those certified as homicides as well as those certified as suicides are significantly less likely to be issued by American medico-legal practitioners (OR = 0.026 for the former and OR = 0.182 for the latter, respectively). This is more than likely due to the drastically different ways in which medico-legal systems are set up in France and in the United States. Although the medico-legal systems in the United States strictly outline types of deaths falling within medical examiners’ or coroners’ jurisdictions, medico-legal investigations in France are performed on an ad hoc basis, on request by the prosecutor, when the physician in charge of determining the death identifies a medico-legal objection to signing the death certificate (Neuilly, 2011). This creates a situation in which only a minority of cases investigated by French medico-legal practitioners is classified as natural or accidental, as opposed to a majority of the American cases.

Second, although gender of medico-legal practitioner was the only coded individual characteristic, it was a significant one. Indeed, compared with accidents, homicide verdicts are more likely to be issued by male medico-legal practitioners (OR = 5.592). The only other systematic attempt at ascertaining differences between medico-legal practitioners is, to our knowledge, Goodin and Hanzlick’s 1997 survey of more than 700 medical examiners and coroners in the United States. Having presented their respondents with vignettes of various contentious death scenarios, the authors examined their

### Table 2. Multinomial Logistic Regression of Manner of Death (N = 459).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Homicide</th>
<th></th>
<th>Suicide</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>95% CI</td>
<td>OR</td>
<td>95% CI</td>
</tr>
<tr>
<td>American examiner</td>
<td>0.026**</td>
<td>[0.003, 0.272]</td>
<td>0.182*</td>
<td>[0.041, 0.810]</td>
</tr>
<tr>
<td>Male examiner</td>
<td>5.592*</td>
<td>[1.159, 26.983]</td>
<td>1.266</td>
<td>[0.516, 3.106]</td>
</tr>
<tr>
<td>Weapon usage</td>
<td>2.843*</td>
<td>[0.959, 8.428]</td>
<td>1.520</td>
<td>[0.659, 3.504]</td>
</tr>
<tr>
<td>Dead in residence</td>
<td>0.480</td>
<td>[0.155, 1.489]</td>
<td>2.175</td>
<td>[0.944, 5.009]</td>
</tr>
<tr>
<td>Dead at scene</td>
<td>0.239*</td>
<td>[0.067, 0.855]</td>
<td>4.101**</td>
<td>[1.608, 10.459]</td>
</tr>
<tr>
<td>Male victim</td>
<td>1.001</td>
<td>[0.277, 3.613]</td>
<td>2.450</td>
<td>[0.939, 6.396]</td>
</tr>
<tr>
<td>Victim age</td>
<td>0.951**</td>
<td>[0.921, 0.982]</td>
<td>0.998</td>
<td>[0.978, 1.019]</td>
</tr>
<tr>
<td>White victim</td>
<td>0.261*</td>
<td>[0.089, 0.768]</td>
<td>1.389</td>
<td>[0.535, 3.604]</td>
</tr>
</tbody>
</table>

Note. OR = odds ratio; CI = confidence interval.
*p < .05. **p < .01. ***p < .001.
sample’s conclusions on manner of death and found a general lack of agreement. They did not, however, look at individual practitioners’ characteristics to isolate potential factors explaining the variations beyond general training information. Our findings thus open the doors to a new avenue for research.

Discussion

Although our overall results are coherent with the extant literature with regard to decedents’ and incidents’ characteristics and their influence on manner of death classification, our analyses provide innovative findings, first by affirming the importance of cultural definitions and related institutional structures on medico-legal practices and thus mortality statistics (Douglas, 1967; Neuilly, 2011; Rockett, 2010; Varnik et al., 2010), and second, by offering a new avenue to research pertaining to the impact of individual medico-legal practitioners’ characteristics on manner of death decision making, thus furthering Goodin and Hanzlick’s (1997) cornerstone results (Hanzlick & Goodin, 1997).

With regard to cultural and institutional factors, further analyses will have to (a) systematize comparative efforts by replicating such studies in various countries, (b) integrate quantitative analyses with in-depth qualitative data, and (c) establish hierarchical influences of culture and institutional structure beyond the country level, both on a more local and larger scale.

With regard to inter-individual differences between medico-legal practitioners, our findings align with current considerations of implicit bias and its impacts on decision making (Carney, Krieger, & Banaji, 2010; Correll et al., 2007) and suggest that we may need to ask different kinds of questions: Are men more likely to see homicide? Are non-White male decedents more likely to be seen as homicide victims? What comes first? The predictors or the manner of death? The intersection of decedents’ and medico-legal practitioners’ characteristics thus have to be taken into consideration. As such, future research should focus on (a) integrating other individual practitioners’ along with decedents’ and incidents’ characteristics in analyses and (b) diversifying the types of methods used to ascertain such inter-individual differences, from surveys and cross-sectional to experimental designs.

Policy Implications

For close to a century, the conclusion to every study on medico-legal practices and the validity and reliability of mortality statistics has been that more standard and systematic training of a more professional body of practitioners was needed (American Medical Association, 1958; Committee on Identifying the
Needs of the Forensic Science Community, 2009; Hanzlick, 2003; National Municipal League, 1951; National Research Council, 1928), and yet, not much seems to have changed (ProPublica, Public Broadcasting System, & National Public Radio, 2011; Tollefsen et al., 2012; Varnik et al., 2010). On one hand, our findings of cultural and institutional differences support such calls for uniformity and standardization, but on the other, our findings pertaining to gender differences question whether uniformity in training outcomes would be possible or even desirable. What would this uniformity be predicated upon? Within our study, would that uniformity align female medical examiners to males or the reverse? Our findings indeed question the very notion of objectivity, and thus in a diverse, multicultural, and ever-shrinking world, is not a call for uniformity doomed to fail because of its hegemonic undertones? Although ensuring better training will always be desirable, we suggest that true random assignment of cases between practitioners within medico-legal institutions, interrater reliability checks within and between institutions, and data adjustments at the aggregate level based on an in-depth understanding of data limitations at the local, regional, national, and international levels may lead to more fruitful outcomes with regard to data reliability and validity.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was partially funded by a Rutgers University School of Criminal Justice Graduate Student Research grant (2005) and a Washington State University New Faculty Seed grant (2013).

References


**Author Biographies**

**Ming-Li Hsieh**, PhD, is a visiting assistant professor in Criminal Justice Program at University of Wisconsin-Eau Claire. Her research includes an examination of gender gap trends, risk assessment instruments, crime control policies and comparative criminal justice. Her recent work has been published in *Criminal Justice Policy Review, International Criminal Justice Review*, and *Feminist Criminology*, among others.

**Melanie-Angela Neuilly**, PhD, Criminal Justice Rutgers University, 2007; PhD, Human Sciences, Université de Rennes, France, 2008, is an associate professor in the Department of Criminal Justice and Criminology at Washington State University. Her research focuses on lethal violence data collection processes in a comparative and international context. Her recent work has been published in *Homicide Studies, the Journal of Interpersonal Violence*, and *Aggression and Violent Behavior*, among others.