The Homicide Atlas in Colombia: Contagion and Under-Registration for Small Areas

B. Piedad Urdinola*
Universidad Nacional de Colombia, Bogotá - Colombia

Francisco Torres Avilés†
Universidad de Santiago de Chile, Santiago - Chile

Jairo Alexander Velasco**
Universidad Nacional de Colombia, Bogotá - Colombia

Abstract

The homicide atlas in Colombia is a visual representation of both expansion and aggravation of the armed internal conflict for the deadly decades of 1990 to 2009. However, mortality under-registration remains an issue in most developing countries, more remarkably when studying particular causes of death on small areas. This document proposes a Bayesian spatial method to identify mortality under-registration in municipalities. Probability maps help to identify under-registered municipalities in Colombia that coincide with the rise of violence at the turn of the century, which is not captured in vital registration systems. It also shows that women suffer of higher under-registration issues than men. Corrected homicide Atlases facilitate interpretation and the proposed methodology proves to be a good source of under-registration identification in small populations.

Keywords: bayesian spatial analysis, small area estimation, homicide atlas, demography of conflict, mortality under-registration.
El Atlas del homicidio en Colombia: contagio y subregistro en áreas pequeñas

Resumen

El atlas del homicidio en Colombia es una representación visual de la expansión y agravamiento del conflicto armado interno en las décadas más mortales: 1990 a 2009. Sin embargo, el subregistro de la mortalidad sigue siendo un problema en la mayoría de los países en desarrollo, más notablemente en el estudio de causas específicas de muerte en áreas pequeñas. Este documento propone un Método Espacial Bayesiano para identificar el subregistro de la mortalidad en los municipios. Los mapas de probabilidad ayudan a identificar los municipios subregistrados en Colombia que coinciden con el aumento de la violencia a principios del siglo XXI, no capturada en sistemas del registro vitales. También muestra que las mujeres sufren mayores problemas de subregistro que los hombres. Los atlas del homicidio corregidos facilitan la interpretación y la metodología propuesta demuestra ser una buena fuente de identificación del subregistro en poblaciones pequeñas.

Palabras clave: análisis espacial bayesiano, atlas de homicidio, demografía de conflicto, estimación de área pequeña, subregistro de mortalidad.

O atlas do homicídio na Colômbia: contágio e sub-registro em áreas pequenas

Resumo

O atlas do homicídio na Colômbia é uma representação visual da expansão e do agravamento do conflito interno armado nas décadas mais mortais: de 1990 a 2009. Contudo, o sub-registro da mortalidade continua sendo um problema na maioria dos países em desenvolvimento, mais notavelmente no estudo de causas específicas de morte em áreas pequenas. Este documento propõe um método espacial bayesiano para identificar o sub-registro da mortalidade nos municípios. Os mapas de probabilidade ajudam a identificar os municípios sub-registrados na Colômbia que coincidem com o aumento da violência no início do século XXI, não capturado em sistemas de registro vitais. Também mostra que as mulheres sofrem maiores problemas de sub-registro do que os homens. Os atlas corrigidos facilitam a interpretação, e a metodologia proposta demonstra ser uma boa fonte de identificação do sub-registro em populações pequenas.

Palavras-chave: análise espacial bayesiana, atlas de homicídio, demografia de conflito, estimativa de área pequena, sub-registro de mortalidade.
Introduction

Mortality or Disease Atlases have been widely used by Epidemiologists and Demographers to study geographic spread of contagious diseases, mortality rates, and incidence or prevalence rates of certain illnesses. These Atlases help in the study of mortality patterns and distribution across time and/or space and, sometimes, in other socio-economic or demographic variables with public health implications (European Commission, Inserm, UFX, and Theme 3 2002; Hansell et al. 2014; Hay et al. 2013; Linard et al. 2010; Magalhães et al. 2013; Noor et al. 2008; Sinka et al. 2011, 2012; Tatem et al. 2006, 2012). However, to build a detailed mortality atlas it is necessary to have a fair quality of mortality data for small areas, which usually suffers from under-registration that tends to be higher in developing nations (U.N. 1983, 2002, 2005). The problem intensifies once mortality rates split by at least one more dimension, namely, cause of death, sex, or age. Over the last decade, Bayesian models have provided an alternative for under-registration correction (Assunção et al. 2005). This document proposes the use of Bayesian estimation models combined with spatial analysis to help identify under-registration of mortality rates of specific causes, by age and sex in small areas, which can easily be interpreted in a mortality atlas.

In particular, non-communicable diseases have been over passed over by the literature, except by for a few mortality Cancer Atlases (Boyle and Smans 2008; CDC 2014; López-Abente et al. 2007; Moreno et al. 2003; Public Health England 2014), the Chilean Atlas of mortality with the 16 main causes of death (Icaza et al. 2013), and an European Union attempt to portray contagious diseases, cancer, circulatory diseases, ischemic heart diseases, cerebrovascular diseases, external causes of death, suicide, transport accidents and alcohol related causes of death (Eurostat 2009). For that reason, this document presents an Atlas of homicide rates for young adults in Colombia during the record peak-decades, from 1990 to 2009, with a twofold purpose. First, to create an Atlas of homicide rates by sex and municipality that allows the descriptive and standard analysis of the Mortality Rates Atlas to identify gender and spatial differences across Colombia and homicide propagation over time. Second, to apply a statistical model of Disease Mapping, from a Bayesian point of view that allows detection of municipalities with low quality data. This methodology will be useful for all other causes of death, but in the particular case of Colombia, it provides visualization of how the conflict escalates and the identification of zones under conflict that may be hidden due to under-registration in vital statistics.

By studying homicides, one can signal regions in Colombia where under-registration is important, perhaps overlooked by the authorities, and provides clean evidence of the usefulness of the hereproposed here methodology to identify under-registration in small areas. In fact, the cause of death with the lowest under-registration levels in the vital statistics systems is homicide (Fajnzylber, Lederman and Loayza 1998).

The statistical analysis used in this paper is in line with the original work developed by Besag (1974) and later formalized by Besag, York and Mollie (1991). It introduced the use of neighborhood information for small areas and is considered the cornerstone for the so-called Disease Mapping. The estimation process follows Bayesian methods, as it allows estimating risks and their stabilized standard deviations, which leads to more realistic credible intervals for inference purposes. Usually results presented in maps enable descriptive analysis (Lawson 2006) and incorporate data for both the area under study and its contiguous neighbors. Neighboring effects could be important for the study of homicides in Colombia, and elsewhere, as violence eruptions in one place can lead to propagation in nearby places. As violence rapidly spread in Colombia, from the last years of the 1990s to the early years of 2000s, there are several municipalities with well-reported areas under conflict with low reported homicide rates, denying a certain level of contagion between neighboring municipalities. This contagion in small areas, particularly in those with low population density, is persistent and intensifies as the conflict escalates with devastating consequences for both males and females. After this introduction, the literature review follows. Section 3 presents an overview of the data and the methodology. Section 4 depicts the results and Section 5 has the final discussion.

Literature Review

Small area populations or small populations refer to any population measured at the sub-national level smaller than provinces, for any country, and excluding megacities. It also refers to conglomerates of particular small populations such as minorities, religious communities or enclaves of forced migrants. Swanson and Tayman (2012) point out the difficulties to find accurate sources, for both bibliography and data, to produce the estimations...
and projections for small area populations. In most cases the only available source for small areas’ information are censuses and vital registration systems. They provide a comprehensive review of all current available methodologies to deal with the most common issues on the topic, which include pure demographic, statistical and, more recently, spatial or geographical methods.

However, prior to the task of projection for small populations, the input data for any chosen model are vital statistics that suffer several measurement issues. Most vital statistics for small areas are either incomplete due to missing data points over time series or suffer large under-registration problems (Wilson and Bell 2011), which in turn would lead to low quality projections. Under-registration of mortality statistics is a common issue in developing countries at all regional levels, including national estimates, which is true for current Latin American data (i.e. Banister and Hill 2004; Bayona and Ruiz 1982; Coale and Kisker 1986; Dechter and Preston 1991; Hill 2003; Jorge et al. 2007; Timaeus 1991; Timaeus, Chackiel and Ruzicka 1996; Urdinola and Queiroz 2013). To overcome under-registration of mortality at a national level, several methods have been developed since the 1960s with a large consensus on their validity in the demographic community (Moultrie et al. 2004). The following section presents the details and limitations to the methodological construction.

Solutions to produce accurate estimates for small populations’ vital statistics include Dual or Multiple System Estimators, Microsimulations, Neural Networks, Social Network Analysis & Spatial Demography (Mathews and Parker 2013; Swanson and Tayman 2012). The problem with most of these methodologies is the need for auxiliary data, which relies on additional micro-level data, such as household surveys or administrative records, which could also be incomplete or hard to access. In the case of Spatial Demography, auxiliary data derives from geographic information systems or extensive satellite imagery. These limitations are precisely what motivated this proposal of a spatial model with a Bayesian approach that relies entirely on the current mortality data at hand, to both describe observed mortality patterns and to signal small areas with under-registration problems. More importantly, it would reduce the search for auxiliary data, so any country with a vital registration system, which suffers of under-registration could apply the methodology proposed here.

To build a Mortality Atlas, it is common to use standardized mortality ratios (SMR). However, the frequentist approach may lead to unstable estimations of SMR, since the associated variations are related to population density and, as a result, municipalities with lower population produce over-estimated variances. To correct for this problem, the Bayesian paradigm provides a better control, as shown by Marshall (1993), Mollié (2000), Pascutto et al. (2000) and Assunção et al. (2005). However, combining spatial demography and Bayesian estimation is rare in the literature and has proven to be useful for the study of small area differences of SMRs in different contexts such as mortality patterns in Italy (Divino, Egidi and Salvatore 2009) or forecast of HIV prevalence in Uganda and Tanzania (Clark, Thomas and Bao 2012).

In particular, Divino, Egidi and Salvatore (2009) differ from our methodological point of view, as they apply a hierarchical spatial model. This helps by pointing out geographic differences in mortality rates across small areas, despite the fact that they can be inter-related and hold persistent over-time geographic information, which influences observed rates. Their objective is to account for such dependences that can affect other small areas SMR. While, Clark et al. (2012) produced a forecast of specific rates by age, sex of HIV prevalence, despite the difficulties to properly quantify HIV contaminated people. Their analysis uses an extension of a mathematical model for contagion and spread in time, by using a Bayesian application with a maximum likelihood estimation method. Their results follow a mathematical model applied to observed data on population counts and HIV prevalence for small areas.

The methodology proposed here accounts for random effects, an unstructured distribution and geographical effects. It controls for the geographical effect by taking neighbor structure into account in order to stabilize the over (or sub) dispersion that can be produced by the surrounding areas (see epidemiological applications in Best, York and Mollié et al. 2005; Lawson 2006; Richardson et al. 2004). The following section presents the details and limitations to the methodological construction.

Data and Methods

Data

The political-administrative internal division of Colombia has remained constant since the creation of the National Constitution in 1991, with 32 Departments (Departamentos) and Bogotá, the capital city, which is not considered a Department, but is classified independently for all purposes since the
second half of the 20th century. Each Department is fractioned in smaller sub-regions called "municipios" or municipalities, as will be called for the remainder of the paper. The number of municipalities varies with total population per square meter in an area, by Constitutional law. In fact, in 1991, Colombia had almost 1,000 municipalities, but by 2013, there were 1,122 official municipalities in Colombia. For this reason, we used geographic coding data for the 730 municipalities that remained geographically consistent for all years during the entire period under study, 1990-2009. Data was taken from the Otlet Library.

The Official National Statistics Bureau in Colombia (Departamento Administrativo Nacional de Estadística - DANE) provided homicide records by age, sex and municipality. In Colombia, overall adults’ death under-registration (ages 15-60) range from 15% to 55%, depending on the methodology and years under study. However, most studies agree that despite these high-medium levels of under-registration, the overall trend is captured (Bayona and Pabón 1983; Bayona and Ruiz 1983; Flórez 1985, 2000; OMS 2001; PAHO 1999, 2003, 2009). This makes the Colombian case an attractive one to apply the here methodology proposed here. As mentioned before, this paper focuses on homicides, because the legal process demands an autopsy for all cases, which makes this particular cause the most consistent of all causes of death in terms of data collection (Fajnzylber, Lederman and Loayza 1998).

The reliability of causes of death has been a cause of debate over decades. Vital registration systems depend heavily on Medical Doctors who may fail to identify the leading cause of death, for several reasons. For instance, death certificates can be filled out without a treating Medical Doctor, as occurs in many developing countries. This debate on quality of causes of death has led to the creation of “verbal autopsy” methodology (WHO 2012), as an instrument that could help identify the main cause of death by interviewing relatives, next of kin and caregivers of the deceased. Thus, using homicides avoids the distraction of misreporting cause of death, while it allows accounting for actual under-registration issues.

Homicide counts from 1990 to 1999 were provided in wide age groups while data from 2000 to 2009 was given in five-years-age-groups. All data was homogenized to the former decade. There were three data points (combination of year and municipality) with missing homicide data, but such municipalities had serial years with homicide records. For those three data points, homicide counts are imputed with the mean number of homicides for neighboring municipalities for the year-municipality they were missing.

Yearly population estimates by age, sex and municipality are calculated from the two latest National Census records (DANE), carried out in 1993 and 2005, by using linear extrapolation of the exponential model of population growth. To maintain consistency in the time series, municipalities for which there was no geographical information in the Otlet library were eliminated.

Both uniform age groups of homicide counts and population estimates allow the calculation of yearly homicide rates by age and sex per municipality for the entire period from 1990 to 2009. Age groups for each sex correspond to ages: 0-4, 5-14, 15-44, 45-64 and 65+. Then, municipal homicide rates by age and sex were standardized by using the total Colombian population as standard population. Standardization allows eliminating population size effects and is used for the remaining calculations. Results focus on adult mortality, between ages of 15 and 44, as child, youth mortality, and elderly mortality suffer from much higher under-registration. Those age groups also have many zero rates of homicide, which requires a different type of estimation model (for instance, zero-inflated distribution). Nevertheless, all estimations are done for each of the age groups described above and combining all ages or total deaths.

Statistical Methodology

This paper follows Bayesian hierarchical methods, proposed by Besag, York and Mollié (1991), to draw data for homicide rates in Colombia from age-grouped structure of time series. There is only one previous attempt at building the homicide atlas in Colombia. Cortés Rueda (2005) portrayed the evolution of homicide and massacre counts for years from 1999 to 2003 at the departmental level and for “cabeceras municipales” or municipal heads. His results show a correlation between economics sources and the increments of violence, understanding as economic resources several agricultural and mineral resources, including cocaine plantations and access to infrastructure. However, that attempt follows a very different methodology in several aspects from the one presented here. First, it does not account for under registration issues; second, it is not portrayed at the municipal level, only at the municipal

1 Available at http://otlet.sims.berkeley.edu/imls/world/shp.files.
heads level; third, it ignores demographic forces by only incorporating counts and not rates, which are not standardized, as they should be by age and sex, two of the basic variables that define mortality and homicide patterns at national and sub national level; finally, it over weights data by spreading the overall effect of departments, rather than using municipal data.

For these reasons, the methodology presented here will produce maps useful for researchers and specialists in crime data or conflict analysts as it provides incidences behavior in small populations in Colombia over long yearly series.

Besag’s approach allows quantifying uncertainty and translating it into confidence intervals over the estimates. It also generates consistent estimates by adding the geographical dimension. Those two conditions enable estimating the geographic distribution of homicide risk, including municipalities, whose statistical size does not allow estimating their risks with reasonable accuracy.

For each year the assumption is that the number of y, cases measured in the i-th municipality, is Poisson-distributed with mean (e, r) where e is the expected number of cases obtained from the indirect standardized method of rates and the unknown quantity r, is the risk associated to each i-th municipality, i=1,...,n. The structure of the model is as follows:

\[ y_i \sim \text{Poisson}(e_i, r_i) \]
\[ \log(r_i) = \alpha + u_i + s_i \]

Where the real quantity, \( \alpha \), represents a constant effect over the complete region \( u_i \) and \( s_i \) represent the unstructured and spatially structured random effects respectively, with \( u_{ij} : i=1,...,n \) identically, independently and normally distributed with zero mean and common variance \( \sigma_u^2 \), and \( s = (s_1,...,s_n) \) are assumed to follow an intrinsic autoregressive random effect (icar), that is:

\[ s_i | \{ s_j : j \neq i \} \sim N(\mu, \sigma_i^2) \]

Where \( \mu_i \) and \( \sigma_i^2 \) are the weighted average and variance incorporating the immediate neighbors. Also,

\[ \mu_i = \frac{\sum w_{ij} s_j}{\sum w_{ij}} \quad \text{and} \quad \sigma_i^2 = \frac{\sigma^2 - \sum w_{ij}}{\sum w_{ij}} \]

While \( \sigma^2 \) is the variance over the entire region and \( w_{ij} \) represents the 0-1 weights, denoting as \( w_{ij}=1 \) when the i-th and j-th municipalities share a common boundary, and \( w_{ij}=0 \) in other case.

All risk estimations used a Bayesian estimation method. Therefore, in order to represent no information over the unknown parameters, non-informative priors for the hyper-parameters \( \sigma_u^2 \) and \( \sigma_i^2 \), both follow an inverse Gamma distribution with parameters 0,001. Finally, a flat prior density (dflat) is assumed for the parameter \( \alpha \), as prior knowledge, which is equivalent to elicit an improper uniform prior. The latter is an important subject to consider; because of this assumption, the sum to zero constraint for the spatial random effects, already implemented in the icar distribution in the Winbugs, works properly for this analysis. The estimation process used the Specialized Bayesian software Openbugs (Lunn et al. 2009), and obtained MCMC chains whenever necessary to make inference. To obtain the results we applied straightforward codes.

For technical aspects, the choropleth maps were built using the posterior median of the standardized homicide rates, after running a chain with 75,000 iterations, burning in the first 5,000 and keeping those resulting after a systematic sampling every 50 iterations. Convergence was checked before making the inference, which was finally made with samples of 1,400 MCMC iterations.

**Results: Homicide Adult Mortality Rates in Colombia from 1990 to 2000**

**Peculiarities of the Colombian Case**

During the second half of the twentieth century, Colombia underwent a series of civilian conflicts that left Colombia with an excess of adult male mortality when compared to other countries in the region, mainly explained by the increase in homicides. Homicide rates, and particularly young-adult (ages 15 to 44) homicide rates, are a good proxy for the intensity of armed conflicts in the world. Homicide atlases, as the one proposed here, provide an instant graphic representation of homicide rates, as well as the escalation of violence in Colombia over time. Moreover, homicide counts suffer much lower under-registration than other causes of death for two main reasons (Fajnzylber, Lederman and Loayza 1998). First, all identified corpses by external causes of death (suicide, accidents and homicides) must follow a particular protocol that includes a compulsory autopsy. Therefore, it is very hard to hide these deaths from the vital record system, and even if their recognition is delayed (i.e. mass graves),
the protocol must still be followed and then the vital registration is corrected. Second, this pathological exam clearly identifies the cause of death, unlike internal causes of death that rely only on the official death certificate filled out by the medical doctor who signed it at the time of the death, who may or not be the treating doctor of the deceased. Finally, cause attribution for internal causes of death suffer intrinsic mis-recording problems, highlighted in the literature, which is aggravated in developing countries (Setel et al. 2005; WHO 2012).

The recent history of Colombia points out the first eruption of the violence of the second half of the twentieth century in the mid-1940, with the confrontation of the followers of the two main political parties, Liberals and Conservatives, in the period called “La Violencia” (The Violence). This latter transformed into a war of bandits and continued until the late 1950s. By the mid 1960s, as in many other Latin American countries, socialist guerrillas erupted in rural areas into what is known as the beginning of the current “Armed Internal Conflict” (1965-present). This conflict has persisted for such a long time and with so many different actors that it is a very dynamic process that needs to be revisited over time. In fact, by the beginning of the 1980s, a group of civilians supports the creation of paramilitary groups as a form of resistance to constant guerrilla attacks in places lacking State protection. Finally, the war on drugs has financially fueled the conflict as both paramilitary and guerrilla groups found the financial support in this illicit, but very profitable business.

The conflict, then, escalated in magnitude, as well as in number of actors and related victims. By 1991, the overall homicide rate for Colombia reached the levels of 1953, the peak of the La violencia period. After 1990, the conflict intensifies and violence levels simply escalated, which in turn reflected in persistent high homicide rates, as opposed to shorter peak periods observed in the past. This persistence lasts for almost 15 years; therefore, our interest focuses on studying this particular period from 1990 to 2009, as shown in figure 1. The decrease in homicide rates, beginning in 2000, is the result of a very aggressive State policy by an intense military offensive, directed by former President Uribe (2000-2008).

In geographic terms, during the period of analysis, 1990 to 2009, outlaw armed groups erupted in municipalities with lack of State presence and with sudden economic booms, such as those produced by the discovery and exploitation of emeralds, gold, petroleum oil, bananas, coca leaf plantations and cocaine processing (Montenegro and Posada 2001). In general, the two main guerrilla groups settled in isolated municipalities whose main economic activity is the exploration and transportation of petroleum oil, coal or gold, production and transportation of bananas and even a few coffee growing regions (Rangel 1999), or wherever there are high fiscal tax rents receipts and these groups can threaten local majors in exchange for such resources (Peñate 1999). On the other hand, paramilitary groups react to political attempts that could empower guerilla groups (Romero 2003) or directly confront guerrilla groups in highly profitable territories, such as those with illicit crops or transportation of illicit drugs (Echandía 1999).

Figure 1. Homicide rates and counts for Colombia, 1938-2009.
Source: authors’s calculations from Annual Information from Colombian Official Statistics Yearbook - DANE.
This behavior rules out the intuitive perception that poverty is directly related to the eruption of violence in the Colombian Armed Conflict, at least since 1990 (Montenegro and Posada 2001; Urdinola 2004). This last fact is very important for this research, as it avoids confusion between poverty and homicide under-registration. That is, usually the poorest regions hold the highest levels of mortality under-registration, but they are not necessarily the regions with actual higher homicide rates. It also reinforces the idea that neighboring municipalities influence each other’s homicide rates, as violence erupts because of a sudden economic boom, usually related to the ownership of an agricultural or mineral product, or the transportation of illicit drugs, which finally places the areas in conflict.

Results

Figure 2 shows the results from the model applied to young adults’ (aged 15 to 44) homicide rate, for the overall period 1990-2009, (as reference, annex 1 presents the official political division of Colombia). This first map serves as guide to read the following results. Panel A, top-right, shows the map with the actual homicide rates. They are presented as the observed rate per municipality over the expected, the frequentist estimate, where the latter is the total expected value for all municipalities. This representation enables visualizing the intensity of homicide rates per municipality in contrast to total national levels. This paper takes advantage of the poetic meaning of black representing death and white peace, as no homicide takes place. Blacked out municipalities are those with the highest homicide rates (above 1.1), much higher than national average, while those in white had the lowest levels (between 0 and 0.3), compared to average national standardized rates. Therefore, gray tones reflect municipality levels that range from higher risk (darker), above national average levels, to those reaching the lowest risk compared to national average levels (lighter).

Panel A, in figure 2, identifies regions with the highest risk of homicide rates that coincide with those signaled by experts, as the most affected by the conflict in Colombia: remote municipalities with lack of State presence and with high income from profitable economic activities. For instance, several municipalities with the highest production of petroleum oil in Colombia located in the Departments of Casanare and Arauca (Southeast departments), or municipalities at the central coffee growing region, or the banana plantations in lower Magdalena. Panel B, bottom left, show the accuracy of applying the methodology proposed here for each municipality, which is only informative and from which no actual inference is made.
Finally, Panel C (right bottom) shows the spatial trend proposed by the model with the estimated risk of standardized homicide rate assigned to each municipality. One of the advantages of this model is that it graphically represents under-registered observations, by simply contrasting Panel A and Panel C. For instance, the municipalities that passed from zero homicides, uncolored municipalities in Panel A, to completed blacked out in Panel C. The results prove that at least 132 municipalities are detected by the model to be extremely under-registered. Moreover, all other municipalities that do not present extreme results, but still show darkening from observed (Panel A) to predicted (Panel C) values, that is, pass from one category to another, for at least a grade darker, add up to a total of 205 municipalities. Not surprisingly, most of such municipalities are located in isolated areas located in the so-called Macarena region and in the Departments of Llanos Orientales (Arauca, Casanare, Vichada and Meta). Areas aside from the three Andean mountain chains, inhabited by nearly 70% of the total population have the lowest population density in Colombia and therefore have usually been under-counted by census and vital registration systems (Urdinola and Herrera 2015).

There is one caveat to these plain results. Panel C shows some municipalities with a lower risk than originally presented in Panel A. There are 47 municipalities whose color faded from Panel A to Panel C. This is the result of the geographic smoothing by the model. As mortality is under-registered but not over-registered, we suggest passing over these results and focusing solely on those municipalities that are spotted as with higher risk than observed data, that is, all those that darkened. Researchers and policy makers could use a final version of the maps with the corresponding observed SMRs for those that faded.

As the objective of the paper is to provide time differences in homicide atlases, consecutive chronological three-year periods between 1990 and 2009 are chosen to present the results for each sex separately. This allows observing time trend differences and the precision of the model for each gender, as intensity of homicides for male and female are different. In order to avoid annual homicide peaks or severe under-registration for a given year, they are smoothed by building rates for three consecutive years. In general, homicide rates for young adults affect, by a wide difference, men more than women, and in a country under severe conflict such as Colombia, this effect is particularly strong. All following figures, from 3 to 12, are interpreted as figure 2, therefore, the analysis will only refer to the contrast of Panel A and C in all cases and, following other demographer’s visual representations (i.e. age-pyramids), which also present males before females.

The first remarkable time trend is the escalation of the conflict over the years affecting both sexes, as portrayed by more municipalities turning into dark in all A-Panels. Second, male and female depict very similar maps. They represent almost the same municipalities with the same intensity for homicide for all sub-periods. In fact, if one were to overlap male and female maps of observed homicide rates, the map is almost the same, with few exceptions. This particular result coincides with the idea that young adult homicide rates are a good proxy for conflict, despite the fact that it aggregates homicides from and outside the conflict (i.e. passionate crimes).

More importantly, the results of the model increase in precision, by having shorter time periods and sexes, allowing under-registration detection for each of such categories. Panels C for figures 3 and 4 represent the model for males and females, respectively for the 1990-1993 period. The most important result here is that each map captures different trends for each sex. Typically, women have lower probabilities of dying at all ages compared to men, and this pattern is exacerbated for homicides, where young males have much higher rates than females. In particular, observed standardized rates for males vary from 66.6 to 158.3; while for females, they vary between 5.82 and 12.6. Likewise, the model produces probabilities for males from 0.35 to 1.00 and for females between 0.27 and 0.96. More importantly, the model detects under-registration issues for 265 municipalities in the case of males and 391 municipalities for females. The model detects, as expected, more cases in less dense municipalities and is particularly true for women. These results replicate what demographers have detected over the years with overall mortality under-registration in most Latin American countries: a sex bias in vital registration that affects women more than men, even in this low frequency cause of death for females (Flórez 2000; Timaeus, Chackiel and Ruzicka 1996). The same pattern repeats for each of the following years, as presented in figures 5 to 12. That is, the model detects under-registration for each sex in low-density small areas, but there are more under-registered municipalities for females than for males. As violence escalates in Colombia, the number of observed homicides...
increases, as well as the number of municipalities with non-zero homicide rates. The model captures both facts and the model adapts to each levels of male and female, without exaggerating the probabilistic estimations for the former as homicide rates escalate by the end of the century. In fact, they take the highest observed rates observed in 1998-2001 for young males, figure 5-Panel A, and the estimated probabilities produced by the model, figure 5-Panel C. The contrast evidences under-registration in the traditional low-density municipalities from eastern Colombian plains, but adds some municipalities at the South-West region located in Departments of Nariño and Putumayo (Southwest), and some at lower Magdalena region, not previously detected. In fact, the conflict began to erupt in those southern parts of Colombia, poorly captured in the 1998-2001 period, properly corrected by the model, and later captured by the vital registration system in the following period 2002-2005 (figure 7-Panel A), and consequently corrected by the model (figure 7-Panel C).

Figure 3. Young adults (ages 15-44) homicide rate in Colombia. Males, 1990-1993. Observed and predicted by model. Source: own authors’s calculations from Vital Statistics in Colombia.

Figure 4. Young adults (ages 15-44) homicide rate in Colombia. Females, 1990-1993. Observed and predicted by model. Source: own authors’s calculations from Vital Statistics in Colombia.
Figure 5. Young adults (ages 15-44) homicide rate in Colombia. Males, 1994-1997. Observed and predicted by model. Source: own authors’s calculations from Vital Statistics in Colombia.

Figure 6. Young adults (ages 15-44) homicide rate in Colombia. Females, 1994-1997. Observed and predicted by model. Source: own authors’s calculations from Vital Statistics in Colombia.

Figure 7. Young adults (ages 15-44) homicide rate in Colombia. Males, 1998-2001. Observed and predicted by model. Source: own authors’s calculations from Vital Statistics in Colombia.
Figure 8. Young adults (ages 15-44) homicide rate in Colombia. Females, 1998-2001. Observed and predicted by model.
Source: own authors’s calculations from Vital Statistics in Colombia.

Figure 9. Young adults (ages 15-44) homicide rate in Colombia. Males, 2002-2005. Observed and predicted by model.
Source: own authors’s calculations from Vital Statistics in Colombia.

Figure 10. Young adults (ages 15-44) homicide rate in Colombia. Females, 2002-2005. Observed and predicted by model.
Source: own authors’s calculations from Vital Statistics in Colombia.
Discussion

This document presents a novel method to detect mortality under-registration for small area data. The methodology proposed here follows the recent literature on Bayesian estimation applied to the estimation of demographic rates, but combined with Spatial Demography, where little has been produced, by exploiting the advantage that small area SMR tends to be correlated for several diseases.

This paper presents the results for the Colombian case of homicides for young adults with a twofold purpose. First, to apply this methodology to a non-contagious disease that, for its nature, can also benefit from the use of spatial correlation, a variable usually passed over in the literature. Second, to test for reliability of the model between larger and smaller death rates, as corresponds to the pattern for males and females, respectively, and variations over time.

The results prove the usefulness of both the construction of the homicide atlas for Colombia and the building of estimated risks of homicide for young adults, which lead to identifying under-registration for a particular cause of death in small areas. First, the model adjusts to observe SMR and does not exaggerate detection at any period. That is, even when violence escalates by the end of the twentieth century, the model does not pass over the highest observed rates.
in each area. Second, less populated areas suffer higher under-registration issues, compared to more populated areas; this coincides with overall regional mortality patterns reported from previous studies, where typically isolated populations from the so-called Antiguos Territorios Nacionales, located outside the Andean mountain chains suffer much higher under-registration problems. Finally, female homicides rates are of worse quality than those of males. This also coincides with previous studies about overall quality mortality patterns for Colombia and helps to point out a much-needed additional effort to avoid correlational and gender bias in the Colombian vital registration systems.

The results contrast to mortality atlases developed for other nations by using the methodology to reduce under registration issues, a pattern only observed in vital statistics of developing nations for under studied causes of death. Also, the results contrast to the construction of disease mapping for Colombia or other nations for other kind of diseases by adding the neighboring information that adds the geographical transmission, which is clear in this case, for homicide rates.

These results are possible by the contrast of the maps with the actual and the predicted risk of homicide rates per municipality. The predicted values by the model include those that should be higher as well as those that are smoothed to lower levels. Since over-registration of death is impossible, the reader must be careful to dismiss those results while doing the interpretation.

The use of the model proposed here is not solely limited to the Colombian case of homicides. This model proves to be useful to signal overall mortality under-registration by small area, the propagation of contagious diseases over time and the study of particular death causes that could have an underlying spatial correlation. In this particular case, the model detects under-counted homicide rates that could help to target overpassed over areas with an incipient outbreak of violence, which could be controlled in time, which is particularly true for the case of undetected female homicides in low-density areas, that could be indicating a phenomenon disregarded by the authorities.

Acknowledgments

The genesis of this work was discussed by the authors during a visit by Torres-Avilés to the Department of Statistics of Universidad Nacional, and was partially funded by Fondecyt Grant 11110119. This work was finished during while Urdinola was a Santo Domingo Visiting Scholar at Harvard University. We also want to thank to the participants at the luncheon Seminar at the Population Center at Harvard University and the helpful comments of two blind referees and the editor of Cuadernos de Geografía.

References


B. Piedad Urdinola
Associate Professor. Department of Statistics. Universidad Nacional de Colombia, Bogotá.

Francisco Torres Avilés
Associate Professor, Department of Computation Mathematics and Science. Universidad de Santiago de Chile (post mortem).

Jairo Alexander Velasco
Statistician, Department of Statistics. Universidad Nacional de Colombia, Bogotá.
Annex 1. Official Colombian Map for the Political Internal Division

Source: Instituto Geográfico Agustín Codazzi (IGAC).