

Los que faltan por contar:  
Una estimación de la violencia letal en Casanare, Colombia

Tamy Guberek, Daniel Guzmán, Megan Price, Kristian Lum y Patrick Ball

Benetech Human Rights Program



16 de febrero de 2010

# 1. Introducción

Casanare es un extenso departamento rural de Colombia que cuenta con 19 municipios y una población de casi 300.000 habitantes ubicados en las faldas de los Andes y los Llanos Orientales.<sup>1</sup> Los diversos actores armados del conflicto colombiano han operado en Casanare, entre ellos los grupos paramilitares, las guerrillas y las fuerzas armadas colombianas. Muchas personas de Casanare han sufrido muertes violentas y desapariciones.

Pero ¿cuántas personas han sido asesinadas o desaparecidas? Por razones de política, de responsabilidad y de esclarecimiento histórico, esta pregunta merece una respuesta válida. El Programa de Derechos Humanos de Benetech utiliza la técnica de Estimación por Sistemas Múltiples (ESM) para estimar el número total de homicidios y desapariciones en Casanare entre 1998 y 2007. El presente informe explica por qué a menudo ha sido difícil calcular con precisión las víctimas de homicidio y desaparición, y por qué es importante asegurarse de que esas personas sean contadas adecuadamente. Después, exponemos una metodología desarrollada para estimar tanto el número de víctimas conocidas como el número de víctimas que nunca han sido contadas.

Cualquier recuento de la violencia letal resultaría incorrecto si asumiéramos que cualquier conjunto de datos individual o combinación de conjuntos de datos contiene un conteo exhaustivo de homicidios y desapariciones. Los registros de hechos violentos que mantienen las instituciones gubernamentales y no gubernamentales contienen algunos, pero no todos, los registros de violencia letal. Las organizaciones que recopilan tales datos pueden no tener acceso a ciertos subconjuntos de una población o a un área geográfica determinada. Algunos informes de hechos violentos pueden resultar más fáciles de ubicar que otros, y los conjuntos de datos que resulten podrían tener un sesgo hacia esos casos. ¿Cómo podemos superar tales dificultades?

Las respuestas correctas sobre el número de homicidios y desapariciones dependen de la estimación estadística para superar los patrones de reportaje complejos e incompletos. Es extremadamente importante llegar a cálculos correctos: las estimaciones apropiadas pueden ayudar a dar razón de las víctimas sin identificar y sin registrar en los recuentos históricos y a orientar el desarrollo de políticas para explicar la violencia del pasado. Por el contrario, con el uso de cifras sesgadas o incompletas se corre el riesgo de perderles la huella a algunas de las víctimas y de generar un trauma continuo para la sociedad. Las víctimas que permanecen sin documentar en ningún conjunto de datos se vuelven invisibles, apartadas no sólo de sus propias vidas y de las vidas de sus seres queridos sino de la memoria histórica.

Puesto que consideramos que los conjuntos individuales de datos son incompletos, a la hora de generar estimaciones estadísticas preferimos usar todas las listas o conjuntos de datos disponibles de homicidios y desapariciones. El procedimiento de estimación que se utilizó para calcular la magnitud de los homicidios y desapariciones en Casanare se llama Estimación por Sistemas Múltiples.<sup>2</sup> El método ESM exige que los analistas revisen cuidadosamente todos los incidentes conocidos en las diferentes listas para determinar si algunos de los casos, ya sea dentro de una misma lista o a través de las listas, se refieren a la misma víctima. Cruzar los casos que figuran en más de una lista permite a los estadísticos modelar el proceso por medio del cual se registran los hechos violentos y estimar el número de casos que no han sido contados. Luego, el método ESM utiliza el número de casos únicos observados en cada una de las listas en combinación con el número de traslapes para estimar el número total de víctimas. El uso de un método científicamente riguroso y transparente para “contar los que faltan por contar” significa que los resultados están menos expuestos a reclamos de parcialidad o sesgo.

El trabajo que se presenta a continuación construye sobre la base del estudio previo sobre los desaparecidos del Casanare.<sup>3</sup> Decidimos continuar nuestra investigación en Casanare por tres razones: 1) La distinción

---

<sup>1</sup>DANE: Censo 2005, Bogotá 2006.

<sup>2</sup>Chandra C. Sekar y Edwards W. Deming. Sobre un método de estimación de tasas de natalidad y mortalidad y el alcance de los registros. *Journal of the American Statistical Association*, 44(245):101-115, 1949.

<sup>3</sup>Guzmán, Daniel, Tamy Guberek, Amelia Hoover, Patrick Ball. 2007. “Los desaparecidos de Casanare”. Programa de

entre homicidios y desapariciones con frecuencia es borrosa. Se presume que algunas personas desaparecidas podrían estar muertas. Para comprender la magnitud de la violencia que afecta a Casanare, decidimos yuxtaponer el análisis de los homicidios y de las desapariciones.<sup>4</sup> En todo el análisis que sigue presentamos los resultados de los homicidios y las desapariciones juntos, de tal manera que los lectores puedan comparar el patrón y la magnitud de cada uno de estos hechos. 2) Desde la publicación de nuestro informe de 2007 sobre los desaparecidos de Casanare, hemos logrado mejoras metodológicas significativas en nuestra aplicación de ESM. Estos avances nos permiten incluir todos los conjuntos de datos disponibles y captar con mayor precisión el rango de incertidumbre en las estimaciones. 3) Incorporamos al análisis nuevos datos que han sido compartidos con el Programa de Derechos Humanos de Benetech desde la publicación del informe de 2007. En el presente estudio hemos utilizado la información sobre las víctimas de homicidios y desapariciones que aportan 15 conjuntos de datos.

Estas 15 diferentes fuentes de datos provienen de agencias estatales – entre ellas instituciones del gobierno, de seguridad, forenses y judiciales – y de organizaciones de la sociedad civil. Con base en estos 15 conjuntos de datos y nuestros avances metodológicos, estimamos que en el período 2000-2007 hubo entre 3,944 y 9,983 muertes violentas y que en el período 1998-2005 se presentaron entre 1,270 y 5,552 desapariciones en Casanare. En la Sección 2 presentamos y discutimos estas estimaciones en mayor detalle. En la Sección 3, describimos los datos reportados y cómo se procesaron para utilizarlos en el análisis. Mostramos también por qué los resúmenes descriptivos de los conjuntos de datos individuales pueden ser engañosos. En la Sección 4, sacamos algunas conclusiones generales. En la Sección 5, esbozamos las áreas que esperamos enfocar en nuestro futuro trabajo.

## 2. Resultados

En esta sección presentamos estimaciones del número total de homicidios y de desapariciones en Casanare entre 1998 y 2007. Puesto que una de nuestras metas en este análisis es calcular las estimaciones por año, nuestro período de estudio se determinó de acuerdo con los años sobre los cuales existía suficiente información para establecer las estimaciones desagregadas. Con “desagregadas” nos referimos a las estimaciones según el período de tiempo, la ubicación geográfica y el tipo de violencia, tan específicos como los datos nos permitan examinar. Para estimar las muertes contamos con suficientes datos para calcular estimaciones para el período 2000-2007; y para las desapariciones, el período 1998-2005.

Los resultados que obtenemos por medio de la Estimación por Sistemas Múltiples son un rango de valores plausibles para la magnitud de víctimas totales. Cada valor específico en este rango tiene una probabilidad asociada de ser el valor correcto dados los datos disponibles. La cifra que presentamos como la “estimación puntual” es el valor esperado del rango de valores plausibles.<sup>5</sup>

Al interpretar estas estimaciones es importante tener presente que estamos extrapolando de lo que se conoce y se observa directamente a lo que no se ha observado y por ende sigue siendo desconocido. Podemos aceptar que algunas víctimas no hayan sido incluidas en la documentación existente; pero sin métodos estadísticos, no tenemos manera de establecer cuántas están indocumentadas. Sin una técnica estadística para estimar un total que va más allá de lo que se observa, no podemos evaluar los argumentos sobre la magnitud total de las víctimas. Usando el método ESM, reducimos el campo de incertidumbre, de no saber nada sobre las víctimas indocumentadas a un rango plausible en que se ubica el total – un rango en el que tenemos un

---

Derechos Humanos de Benetech. <http://hrdag.org/resources/publications.shtml>

<sup>4</sup>Hemos incluido los homicidios y las desapariciones de acuerdo con la definición más amplia que utilizan los conjuntos de datos que se nos proporcionaron. En materia de muertes, se incluyen todos los homicidios (la acción de un ser humano que asesina intencionalmente a otro ser humano) excepto el suicidio y las muertes accidentales; en cuanto a las desapariciones, se incluyen las personas desaparecidas cuyo paradero se desconoce. La manera ligeramente diferente como cada fuente define estos hechos letales se describe en la Sección 3.1.

<sup>5</sup>El valor esperado es el promedio de todos los valores en el rango ponderado por la probabilidad de cada valor.

alto nivel de confianza. Este rango de incertidumbre se expresa por los intervalos de credibilidad que están alrededor de la estimación puntual.

Los “intervalos de credibilidad” se basan en una distribución completamente especificada, lo que significa que estamos seguros en un 95 % que el número correcto de hechos violentos está dentro del intervalo (dado que el intervalo se basa en nuestros datos observados). Esta es una interpretación más directa de la incertidumbre que la de los intervalos de confianza tradicionales. Los “intervalos de confianza” tradicionales nos dicen que si realizáramos el estudio 100 veces, 95 de las veces esperaríamos que la respuesta cayera dentro del rango reportado.

Los intervalos de credibilidad de 95 % reportados en este análisis pueden parecer amplios; pero es importante tener presente que captan la incertidumbre relacionada con muchos y distintos conjuntos de datos *y* con el uso de los múltiples conjuntos de modelos estadísticos necesarios para calcular la estimación. En este proyecto hemos adoptado un enfoque abierto frente a la información disponible. Intencionalmente hemos evitado preferir algunos conjuntos de datos a otros, pues escoger conjuntos de datos equivaldría a asumir opciones políticas sobre cuáles de ellos son “mejores”. Las diferentes combinaciones de conjuntos de datos producen estimaciones ligeramente diferentes. Hemos decidido incluir todas las combinaciones de conjuntos de datos y examinar la distribución de las estimaciones que resultan para evitar asumir qué conjuntos de datos podrían ser los mejores. La consecuencia estadística de nuestra decisión es que la incertidumbre que rodea las estimaciones (medida por los intervalos de credibilidad) es sustancial. Creemos que estos intervalos de credibilidad dan un reflejo correcto de la incertidumbre real en torno a las estimaciones del número total de hechos letales en Casanare.<sup>6</sup>

## 2.1. Estratificación

Reconociendo que Casanare es diverso de muchas maneras, estratificamos el departamento en subsecciones y subperíodos más reducidos e internamente homogéneos. Estratificamos por tipo de violencia (muerte o desaparición), por espacio (grupos de municipios) y por tiempo (año). Hay dos razones complementarias para estratificar: la primera es satisfacer uno de los supuestos metodológicos del ESM, que cada caso en la población tiene la misma probabilidad de ser incluido en cualquier conjunto de datos (ver el Apéndice). La segunda es que desde una perspectiva de investigación, una estimación global que abarca todo el área de estudio puede esconder diferencias en los patrones y la magnitud de los homicidios y las desapariciones. Por el contrario, las estimaciones estratificadas nos permiten una comprensión más precisa de las diferencias entre períodos y regiones.

Aunque consideramos tanto las desapariciones como los homicidios violencia letal, reconocemos que el método que se utiliza para perpetrar los dos actos varía; y, además, que los dos se reportan de manera diferente. Por ello realizamos los cálculos necesarios separadamente para cada uno de los dos tipos de violencia.

En cuanto a la estratificación espacial que divide los hechos violentos según el lugar donde se realizaron, no tuvimos los suficientes datos para calcular una estimación separada para cada municipio. Agrupamos los 19 municipios de Casanare por similitud en su experiencia en el conflicto, basados en la orientación que recibimos de expertos locales. También agrupamos los datos de tal manera que los municipios en un grupo sean contiguos.

Se agruparon los siguientes municipios:

---

<sup>6</sup>Sería posible, en algún proyecto futuro, hacer suposiciones sobre cuáles conjuntos de datos son los mejores y usar sólo esos en las estimaciones. Al utilizar un número menor de conjuntos de datos la incertidumbre en torno a las estimaciones sería menor que lo que describimos aquí. Queremos ser explícitos en afirmar que lo uno compensa lo otro: un supuesto adicional reduciría la incertidumbre. En este proyecto optamos por no entrar a hacer tal suposición.

- Región Geográfica *D - Centro*: Yopal y Aguazul
- Región Geográfica *E - Piedemonte*: Sacama, La Salina, Tamara, Recetor, Chameza y Nunchia
- Región Geográfica *F - Sur*: Tauramena, Monterrey, Villanueva, Mani y Sabanalarga
- Región Geográfica *G - Llanos*: Hato Corozal, Paz de Ariporo, Pore, San Luis de Palenque, Trinidad y Orocué

## 2.2. Estimaciones

### 2.2.1. Gráficas de barras

Presentamos los resultados primero usando gráficas de barras. La sección inferior de las barras en cada una de las siguientes gráficas representa el número de víctimas observado en ese año. Encima de esta barra aparece la cantidad adicional del subregistro, de tal manera que la parte superior de cada barra indica el número total estimado de víctimas. El subregistro se refiere al número de homicidios o desapariciones no observados ni registrados en ninguna lista. Las líneas (o “bigotes”) indican el intervalo de credibilidad de 95 % para cada estimación específica.

Los espacios en blanco en todas las gráficas indican los años para los cuales no hubo suficientes datos para generar una estimación. Los espacios en blanco no deben interpretarse como una ausencia de violencia ese año y/o región. Ver mayores detalles sobre el sentido y las implicaciones de “no tener suficientes datos” en el Apéndice: Metodología.

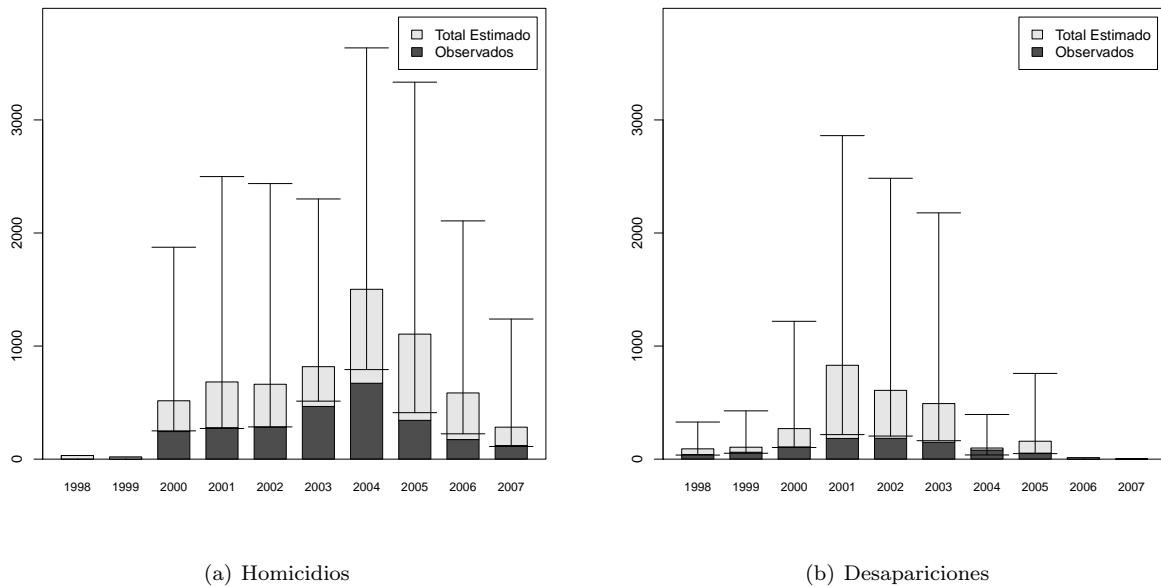


Figura 1: Homicidios y Desapariciones Observadas y Estimadas, por Año

La Figura 1(a) muestra el total de homicidios en Casanare (observados y estimados) en el período 2000-2007. Con base en la estimación puntual parece haberse registrado un aumento gradual en los homicidios

cuyo punto más alto es en 2004, y después una disminución notable. (Debe recordarse que la estimación puntual es el valor esperado del rango de estimaciones plausibles.) Sin embargo, observamos también que los intervalos de credibilidad son muy amplios y se traslapan a través de los años. Esto significa que aunque hay un patrón perceptible entre las estimaciones puntuales, no podemos concluir si esas diferencias entre los años en realidad existen debido a la incertidumbre inherente a nuestras estimaciones.

El patrón en las estimaciones del total de los homicidios sigue el patrón de los datos reportados (parte inferior de la barra). Por ejemplo, el punto máximo en la estimación de homicidios de 2004 es consistente con la historia que relatan los datos subyacentes, que también indican un pico en los homicidios ese año. Lo interesante es que nuestra estimación del subregistro (el número de víctimas no observadas) también es más elevado en 2004. El método que usamos para analizar los datos reportados sobre las víctimas nos permite estimar un número aún mayor de víctimas desconocidas que nunca han sido contadas. Desde luego, aún hay mucho que no sabemos sobre estas víctimas.

El año 2007 es aquel sobre el cual tenemos la menor incertidumbre – es decir, el intervalo de credibilidad más estrecho. El hecho que a las estimaciones de 2004 las rodee una mayor incertidumbre podría ser el efecto del cambio de patrones de documentación. Aunque tenemos más datos para 2004, la proporción de todas las muertes que se documentaron en 2004 podría ser menor que la fracción de todas las muertes documentadas en 2007. Es posible también que la naturaleza de las muertes de 2007 no haya sido igual a la de las de 2004, y que una fracción mayor de las muertes de 2007 (con relación a 2004) caben dentro de los criterios de inclusión de un mayor número de los conjuntos de datos que usamos (ver la Sección 3 para mayores detalles sobre los criterios de inclusión y exclusión para cada conjunto de datos).

En la Figura 1(b), vemos el número de las desapariciones observadas y estimadas por año. Los valores para las desapariciones observadas en 2001-2003 son similares, aunque las estimaciones puntuales (los valores esperados) parecen ser diferentes: la estimación puntual es más alta en 2001. Sin embargo, el alto grado de incertidumbre, medido por los intervalos de credibilidad de 95 %, impiden que concluyamos que la estimación para 2001 es significativamente distinta a las estimaciones para 2002 y 2003. Vale anotar también que hay más incertidumbre, i.e. intervalos de credibilidad más amplios, en las desapariciones que en los homicidios.

Si comparamos las estimaciones del total de homicidios con estimaciones del total de desapariciones, observamos patrones diferentes. Aunque las desapariciones y los homicidios fueron aproximadamente los mismos en 2001, el total de homicidios subió mientras que el de las desapariciones probablemente bajó.

En las Figuras 2–5, presentamos estimaciones desagregadas por homicidios/desapariciones, año y región geográfica. Hemos incluido gráficas para cada región, aunque sólo hacemos observaciones sobre algunas de las gráficas, abajo.

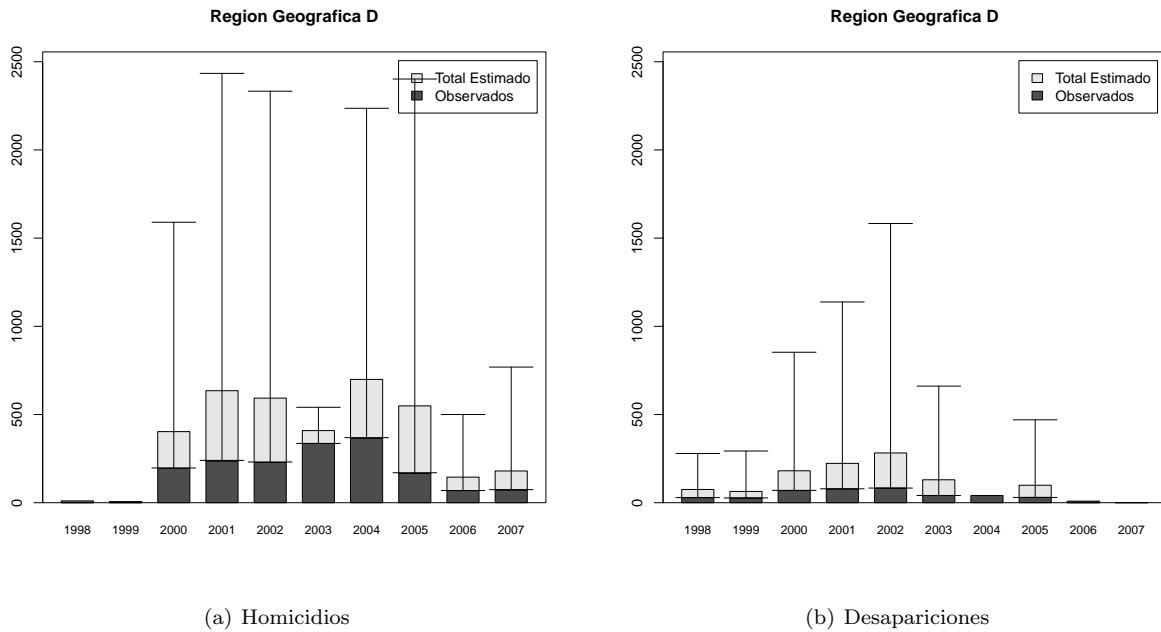


Figura 2: Homicidios y Desapariciones Observadas y Estimadas, por Año para *Centro*

En la región que hemos llamado *Centro*, compuesta por Aguazul y Yopal, estimamos niveles de violencia altos pero constantes. Otras regiones (que se presentan más adelante) tienen patrones de violencia más variables.

En la Figura 2(a), estimamos niveles altos y sostenidos de homicidios, aunque con estimaciones e intervalos de credibilidad más reducidos en 2003 y 2006-07. En la Figura 2(b), estimamos un número de desapariciones relativamente constante para *Centro*, con cifras más bajas que las de los homicidios en la misma región.

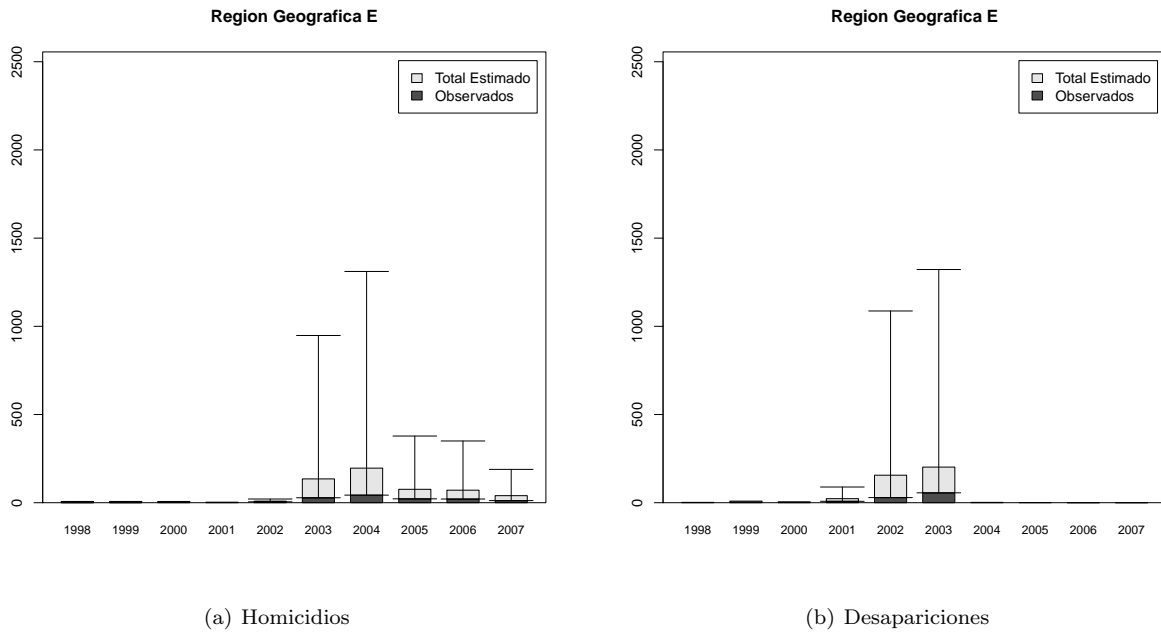


Figura 3: Homicidios y Desapariciones Observadas y Estimadas, por Año para *Piedemonte*

En las Figuras 3(a) and 3(b), estimamos totales relativamente más bajos para la región que llamamos *Piedemonte* que para otras regiones. Sin embargo, hay dos observaciones interesantes: sólo podemos hacer estimaciones de ambos tipos de violencia, homicidios y desapariciones, en 2003; antes de esa fecha, no hubo suficientes reportes para realizar estimaciones de desapariciones, y después de 2003, sólo hubo suficientes reportes para estimar los homicidios. Emerge un patrón interesante cuando calculamos una tasa per capita de violencia dividiendo las cifras estimadas por el tamaño de la población en cada región, abajo (ver Figura 7(b)).



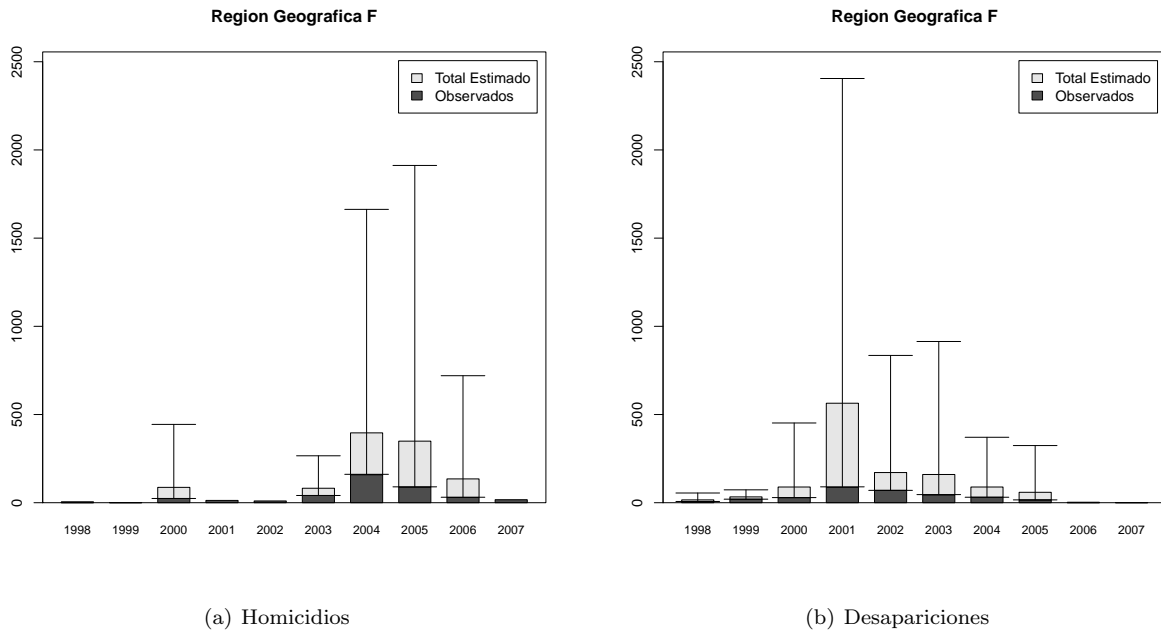


Figura 4: Homicidios y Desapariciones Observadas y Estimadas, por Año para *Sur*

Las Figuras 4(a) y 4(b) muestran las estimaciones para la región *Sur* de Casanare. Estimamos un pico en las desapariciones en 2001. Al comparar las estimaciones de las desapariciones (Figura 4(b)) con las de homicidios (Figura 4(a)) en la misma región, emerge una tendencia. En 2001-2002, hubo pocos reportes de homicidios y por ello no fue posible realizar estimaciones. Al mismo tiempo, las desapariciones registradas y estimadas estaban en su punto más alto para la región – y para todo el departamento. Está claro que los patrones de las desapariciones son sustancialmente diferentes de los patrones de los homicidios, de región en región y de año en año. En 2001 los perpetradores pueden haber estado desapareciendo a las personas en lugar de asesinarlas directamente; en otros períodos, la preferencia se invierte, con más homicidios y menos desapariciones.

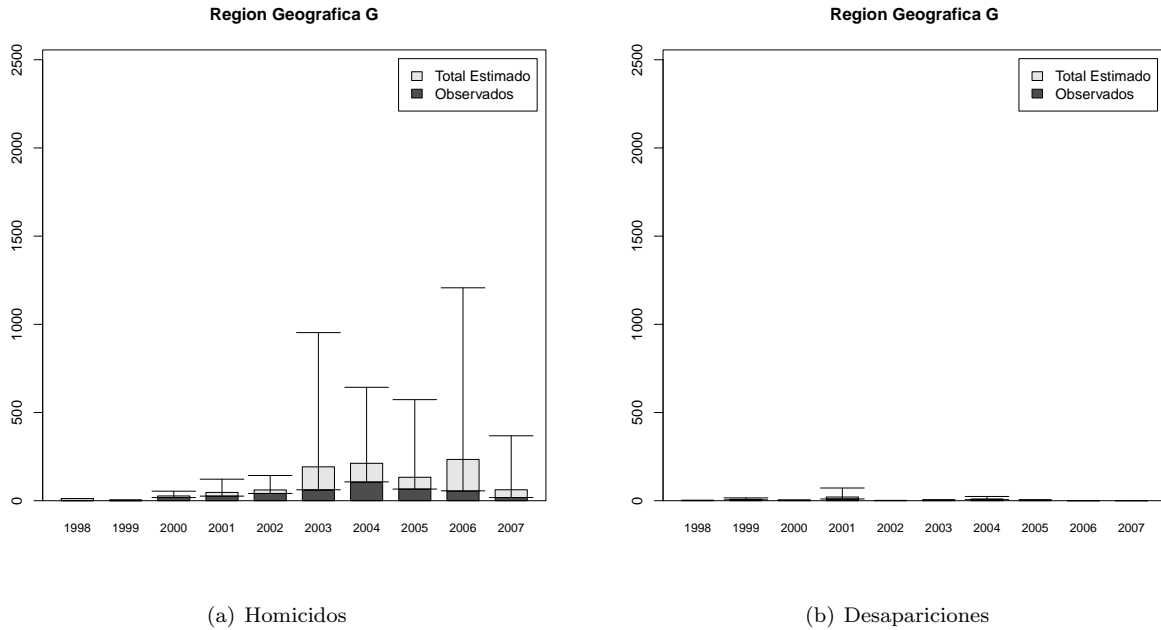


Figura 5: Homicidios y Desapariciones Observadas y Estimadas, por Año para *Llanos*

En la región que hemos llamado *Llanos*, estimamos una cifra baja y constante de homicidios (Figura 5(a)). Sobre las desapariciones fueron pocas las estimaciones que se pudieron realizar (Figure 5(b)). De nuevo, es importante señalar que la ausencia de estimaciones no significa necesariamente ausencia de violencia. La ausencia de estimaciones significa que sabemos muy poco sobre el fenómeno de las desapariciones en esa región.

Estos valores aparecen también en la Tabla 1. Los resultados se discuten en mayor detalle en la Sección 2.3.

### 2.2.2. Mapas

Los mapas que se presentan en esta sección utilizan las mismas estimaciones presentadas en las Figuras 2(a) a 5(b), aunque para facilitar la lectura sólo presentamos la estimación puntual (el valor esperado). Las estimaciones puntuales y los intervalos de credibilidad completos se encuentran en la Tabla 1.

Los matices de color indican un rango para los totales de homicidios o desapariciones por región geográfica al interior de Casanare, lo que le permite al lector comparar la intensidad de la violencia. Las secciones en color gris indican que no contamos con suficientes datos para generar estimaciones. Cabe señalar que los años que se presentan en la Figura 6(a) en homicidios son 2000-2007, y que en la Figura 6(b) los años que se analizan para desapariciones son 1998-2005.

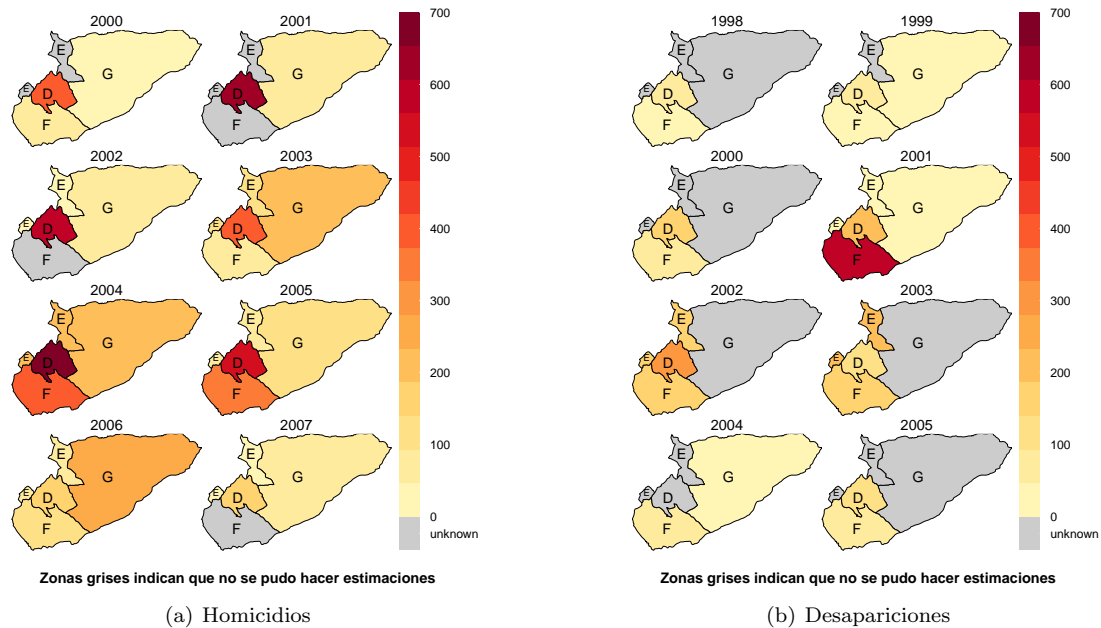


Figura 6: Homicidios y Desapariciones Observadas y Estimadas, por Región y Año en Casanare

En la Figura 6(a) observamos que los números más elevados de homicidios se dieron en la región geográfica *Centro*. Sin embargo, también sabemos que *Centro* tiene la más alta densidad de población. Por ello, abajo presentamos estimaciones como tasas de población, en que dividimos las estimaciones puntuales del total de víctimas por el total de la población por año.<sup>7</sup>

Como vemos en la Figura 6(b), el lugar y el período más frecuente en materia de desapariciones fueron el *Sur* de Casanare en 2001. El único año para el cual pudimos producir estimaciones de desapariciones para todo el departamento es 2001. Para las regiones del *Piedemonte* y los *Llanos*, sólo pudimos calcular estimaciones para 3 de los 8 años.

<sup>7</sup>Utilizamos las proyecciones para el total de la población de Casanare por año calculadas por el DANE. Ver ([http://www.dane.gov.co/daneweb\\_V09/index.php?option=com\\_content&view=article&id=238&Itemid=121](http://www.dane.gov.co/daneweb_V09/index.php?option=com_content&view=article&id=238&Itemid=121)). Debido a que no encontramos las proyecciones anuales desagregadas por municipio, se asignó la misma proporción por municipio que la del censo de 2005.

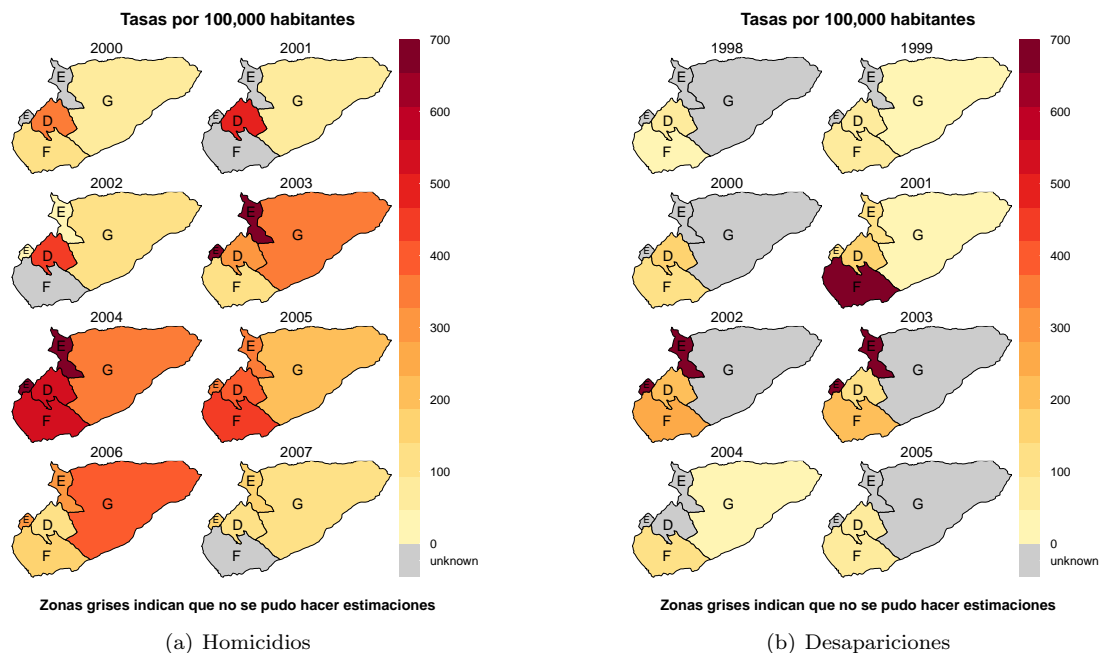


Figura 7: Estimaciones Ajustadas por Población para Homicidios y Desapariciones, por Región y Año en Casanare

La Figura 7(a) ilustra un patrón levemente diferente de homicidios en la región geográfica G (*Llanos*), que no es inmediatamente evidente en la Figura 6(a). Observando la totalidad de Casanare durante todo el período, la región geográfica E (*Piedemonte*) tuvo la tasa más alta de homicidios en 2003 y 2004. En 2006, la región G tuvo la tasa más alta para su población.

La Figura 7(b) confirma que el pico de desapariciones en la región F fue en 2001, como se indica en la Figura 6(b). Sin embargo, lo que no es obvio en el mapa precedente es que la tasa de desapariciones de 2002 y 2003 en la región de *Piedemonte* está entre las más elevadas.

### 2.2.3. Tabla

El Cuadro 1 abajo presenta los datos utilizados en todas las gráficas y los mapas en las secciones precedentes. Por cada año y región, presenta los conteos de las víctimas únicas observadas, los totales estimados y los intervalos de credibilidad correspondientes, primero para los homicidios y luego para las desapariciones.

Cuadro 1: Muertes violentas y Desapariciones

Región Geográfica D - <i>Centro</i>						
Año	Muertes			Desapariciones		
	Observados	Estimados	Intervalo de Credibilidad 95 %	Observados	Estimados	Intervalo de Credibilidad 95 %
1998	9	-	(NA , NA)	29	75	(29 , 279)
1999	6	11	(6 , 29)	27	64	(27 , 293)
2000	195	403	(197 , 1590)	69	181	(70 , 853)
2001	236	635	(240 , 2434)	78	223	(79 , 1138)
2002	228	593	(231 , 2333)	83	282	(83 , 1583)
2003	336	409	(336 , 541)	41	130	(41 , 661)
2004	365	699	(369 , 2236)	41	-	(NA , NA)
2005	166	549	(170 , 2402)	30	99	(30 , 470)
2006	68	145	(69 , 500)	8	-	(NA , NA)
2007	73	180	(74 , 769)	0	-	(NA , NA)
<b>Total</b>	<b>1,682</b>	<b>3,624</b>		<b>406</b>	<b>1,054</b>	
Región Geográfica E - <i>Piedemonte</i>						
Año	Muertes			Desapariciones		
	Observados	Estimados	Intervalo de Credibilidad 95 %	Observados	Estimados	Intervalo de Credibilidad 95 %
1998	7	-	(NA , NA)	2	-	(NA , NA)
1999	7	23	(7 , 113)	9	-	(NA , NA)
2000	7	-	(NA , NA)	5	-	(NA , NA)
2001	3	-	(NA , NA)	8	23	(8 , 89)
2002	5	9	(5 , 21)	29	156	(29 , 1087)
2003	28	135	(28 , 948)	56	202	(56 , 1322)
2004	43	196	(43 , 1311)	2	-	(NA , NA)
2005	22	76	(22 , 378)	1	-	(NA , NA)
2006	21	71	(21 , 350)	1	-	(NA , NA)
2007	12	40	(12 , 189)	0	-	(NA , NA)
<b>Total</b>	<b>155</b>	<b>550</b>		<b>113</b>	<b>381</b>	
Región Geográfica F - <i>Sur</i>						
Año	Muertes			Desapariciones		
	Observados	Estimados	Intervalo de Credibilidad 95 %	Observados	Estimados	Intervalo de Credibilidad 95 %
1998	4	-	(NA , NA)	7	16	(7 , 55)
1999	1	-	(NA , NA)	20	33	(20 , 73)
2000	24	87	(24 , 444)	29	89	(29 , 452)
2001	13	-	(NA , NA)	87	564	(90 , 2405)
2002	10	-	(NA , NA)	70	171	(70 , 835)
2003	41	82	(41 , 266)	46	160	(46 , 914)
2004	159	396	(161 , 1663)	31	89	(31 , 371)
2005	89	349	(90 , 1912)	16	59	(16 , 324)
2006	30	135	(31 , 720)	2	-	(NA , NA)
2007	17	-	(NA , NA)	1	-	(NA , NA)
<b>Total</b>	<b>388</b>	<b>1,049</b>		<b>309</b>	<b>1,181</b>	
Región Geográfica G - <i>Llanos</i>						
Año	Muertes			Desapariciones		
	Observados	Estimados	Intervalo de Credibilidad 95 %	Observados	Estimados	Intervalo de Credibilidad 95 %
1998	12	26	(12 , 88)	3	-	(NA , NA)
1999	4	-	(NA , NA)	5	8	(5 , 16)
2000	18	27	(18 , 54)	5	-	(NA , NA)
2001	26	47	(26 , 122)	10	21	(10 , 72)
2002	41	61	(41 , 143)	2	-	(NA , NA)
2003	61	192	(62 , 953)	6	-	(NA , NA)
2004	104	212	(107 , 643)	6	10	(6 , 24)
2005	66	133	(66 , 573)	6	-	(NA , NA)
2006	54	234	(56 , 1207)	1	-	(NA , NA)
2007	18	62	(18 , 368)	0	-	(NA , NA)
<b>Total</b>	<b>404</b>	<b>994</b>		<b>44</b>	<b>39</b>	

Vale recordar que la falta de estimación no implica ausencia de violencia en un estrato particular. En el caso de la región geográfica G (*Llanos*), las estimaciones faltantes generan un total un tanto extraño a través de los años. Puesto que sólo podemos calcular una estimación del número de desapariciones para tres años en esta región, la suma de estas estimaciones da como resultado un número inferior al del total de desapariciones observadas a través del mismo período de tiempo. Esto se debe al hecho que había registros

de desapariciones en 1998, 2000, 2002, 2003, 2005 y 2006, pero sin embargo en ninguno de estos años fueron suficientes los datos para calcular una estimación.

#### **2.2.4. Comparación con el informe de 2007**

Resulta útil comparar las estimaciones que se presentan en este documento con las de nuestro informe de 2007. Las estimaciones en los dos estudios se calcularon utilizando diferentes conjuntos de datos, diferentes estratificaciones y un método analítico diferente.

Pudimos calcular estimaciones para una ventana de tiempo diferente entre los dos estudios. Las únicas secciones de nuestro informe de 2007 que coinciden con el presente estudio son las estimaciones globales de la totalidad de las desapariciones en todo Casanare en los períodos 2001–2004 y 1998–2005. En nuestro anterior estudio, la estimación puntual para todas las desapariciones en el primer período (2001–2004) fue 821 dentro de un intervalo de confianza de 95 % de (713, 929). Nuestros actuales métodos estiman el número de desapariciones en todo Casanare para este mismo período en 2,028 dentro de un intervalo de credibilidad de 95 % de (840, 4,700). Por un lado, estas estimaciones son conceptualmente diferentes y cada estimación puntual cae por fuera del intervalo de confianza o de credibilidad del otro estudio. Por otro, el intervalo de credibilidad calculado para el presente informe (840, 4,700) es muy amplio y se traslapa con el intervalo de confianza calculado en el informe de 2007.

Para el período más largo, 1998–2005, nuestro anterior informe estimó el número de desapariciones en todo Casanare en 1.239 dentro de un intervalo de confianza de 95 % de (1.088, 1.390). Nuestro actual informe estima el número de desaparecidos para el mismo período en 2,653 dentro de un intervalo de credibilidad de 95 % de (1,270, 5,552). De nuevo, las estimaciones puntuales parecen ser bastante distintas, pero los intervalos se traslapan.

Concluimos que aunque hemos recolectado datos adicionales, hay más desapariciones indocumentadas en todo Casanare que las que habíamos estimado originalmente. Como lo demuestra el amplio intervalo de credibilidad que acompaña esta estimación, queda mucho por saber sobre las desapariciones durante este período de tiempo. Creemos que subestimamos la incertidumbre de la estimación en los cálculos de 2007 y por ende los intervalos de confianza de 95 % son demasiado estrechos. A pesar del tamaño de los intervalos de credibilidad en nuestro actual informe, creemos que reflejan un cálculo más preciso de la incertidumbre de los datos.

#### **2.2.5. Total de hechos letales**

Como se señaló en la sección anterior, la estimación general del número total de desapariciones para todo Casanare entre 1998 y 2005 es de 2,653 al interior de un intervalo de 95 % de credibilidad de (1,270, 5,552). El número total observado de desapariciones para todo Casanare durante el mismo período es de 872, lo que implica que es probable que entre 398 y 4.680 desapariciones no hayan sido reportadas ante ninguna de las 15 fuentes de datos utilizadas en este informe.

Para el período entre 2000 y 2007 estimamos un total de 6,215 homicidios en todo Casanare, dentro de un intervalo de credibilidad de 95 % de (3,944, 9,983). Durante el mismo período de tiempo, se reportaron 2.629 homicidios, lo que implica que es probable que entre 1.297 y 7.336 homicidios no hayan sido reportados a ninguna de las 15 fuentes de datos que se utilizaron en este informe.

En la Sección 3.1, esbozamos algunos de los retos que conlleva la recolección de datos, lo que puede explicar por qué tantos homicidios y desapariciones no se reportan nunca.

## 2.3. Discusión

Nuestra conclusión central es que, aun tras los esfuerzos sustanciales de recolección de datos, persiste el subregistro generalizado de homicidios y desapariciones. Todavía existe una considerable incertidumbre sobre los totales exactos. A pesar de la incertidumbre, hemos podido esbozar ciertos patrones con base en el tiempo, el espacio y el tipo de violencia. El pico de 2001 en las desapariciones generales ocurrió en el *Sur*, precisamente en un momento en que los homicidios reportados en la misma región estaba en su punto más bajo. Los resultados muestran que los homicidios han sido generalizados en Casanare, y alcanzaron su punto más alto en 2004. Este importante número de homicidios en 2004 coincidió a través de todas las regiones de Casanare excepto los *Llanos*, donde estimamos un número elevado de homicidios en 2006. Además, en las regiones en las que los totales estimados parecían relativamente bajos (e.g. E y G) los mapas ajustados por población revelan un efecto más alto en la población. En particular, la región de *Piedemonte* mostró una de las tasas de desaparición más altas del departamento proporcionalmente al tamaño de la población.

Al observar los patrones y la magnitud de los homicidios y de las desapariciones lado a lado, surgen preguntas sobre las estrategias de los perpetradores y la dinámica del conflicto. ¿Por qué varían las desapariciones a la inversa de los homicidios, como se vio en particular en las regiones del *Piedemonte* y el *Sur*? ¿Por qué tendemos a ver el mayor número de desapariciones cuando y donde vemos el menor número de homicidios, como en el *Sur* en 2001 y en el *Piedemonte* en 2002? Y viceversa, ¿por qué hay tan pocas desapariciones reportadas en el *Centro* cuando los homicidios están probablemente en su punto más alto? Para los investigadores de otras disciplinas podría resultar útil combinar las preguntas que emergen de los patrones cuantitativos con sus conocimientos cualitativos y así obtener una mejor comprensión de la violencia letal en Casanare a lo largo de la última década.

La estimación del número de homicidios y desapariciones no observados revela también las dimensiones invisibles de la violencia letal en Casanare. Como vimos en las estimaciones de los totales, es probable que se hayan presentado 70% más homicidios en Casanare que los reportados en todas las listas juntas en el período 2000-2007. Y es probable que se hayan presentado más del doble de las desapariciones reportadas para el departamento, oscilando entre 1,270 y 5,552 en el período entre 1998 y 2005.

Mirando hacia adelante, recomendamos que todas las organizaciones de Casanare hagan los mayores esfuerzos posibles por recopilar datos adicionales sobre la violencia letal. Una mayor tasa de cobertura mejorará la capacidad de los estadísticos y científicos sociales de realizar estimaciones más precisas; es decir, de reducir la incertidumbre en torno a esas estimaciones. Desde luego, estos esfuerzos de recolección de información mejorarán en la medida en que las familias de las víctimas y la sociedad civil en general adquieran mayor confianza y seguridad frente al proceso de reportaje, lo que es más probable si no hay conflicto.

Resaltamos que la tasa de víctimas no documentadas de violencia letal encontrada en Casanare no refleja necesariamente la tasa que podría encontrarse en toda Colombia si los datos estuvieran disponibles. Este análisis demuestra que ningún conjunto de datos por sí solo registra con precisión el número total de personas que probablemente han sido asesinadas en Casanare. Exploramos este tema con más profundidad en la Sección 3.

## 3. Datos

### 3.1. Fuentes de datos

Es difícil recaudar información sobre la violencia. Implica correr riesgos de seguridad, generar confianza e invertir recursos logísticos, financieros y humanos. Muchas instituciones valientes y trabajadoras incursionan en este trabajo a pesar de los retos que ello implica. Es natural que en un lugar tan violento como Casanare algunos grupos se enteren de algunas víctimas y otros grupos de otras. Algunos grupos se enterarán de las mismas víctimas, y algunas víctimas permanecerán por fuera de cualquier esfuerzo de recolección de datos.

Para comprender mejor los datos utilizados en el presente estudio, expondremos brevemente a continuación el alcance de los esfuerzos de recolección de datos de cada organización, de acuerdo con su mandato, que aportó a este estudio.

Todos los conjuntos de datos recaudados por estos grupos fueron compartidos con Benetech o con nuestra coparte, EQUITAS, para su uso en el análisis cuantitativo de datos.<sup>8</sup> Se compartió la información cruda; todo el análisis y las conclusiones en el presente informe son de la autoría exclusiva del Grupo de Análisis de Datos de Derechos Humanos de Benetech y no reflejan la opinión de las organizaciones que recopilaron la información en un principio.

Cuadro 2: Resumen de las Fuentes

Nombre de la Organización	Tipo de violencias	Periodo
INMLCF - Muertes violentas	Muerte	2000-2007
Fiscalía General de la Nación - Desaparecidos	Desaparición	1998-2005
Observatorio de Derechos Humanos de la Vicepresidencia	Muerte	2004-2007
Policía Nacional (DIJIN)	Muerte	2003-2006
CINEP	Muerte y Desaparición	1998-2007
Fondelibertad	Desaparición	1998-2005
Comisión Colombiana de Juristas	Muerte y Desaparición	1998-2007
Coordinación Colombia Europa Estados Unidos	Muerte	2005-2007
INMLCF - Desaparecidos	Desaparición	1998-2005
Fiscalía Seccional Santa Rosa de Viterbo - Desaparecidos	Desaparición	1998-2005
Organizaciones de Familiares de Víctimas	Desaparición	1998-2003
Gaula	Desaparición	1998-2005
Cuerpo Técnico de Investigaciones de la Fiscalía	Desaparición	2003-2005
Equitas	Desaparición	2000-2005
País Libre	Desaparición	1998-2004

Cada institución recopila datos de acuerdo con su mandato institucional. Algunos grupos, como la Comisión Colombiana de Juristas (CCJ), monitorean las violaciones de los derechos humanos. Otros grupos, como EQUITAS, documentan las víctimas para efectos de su identificación humanitaria. El *Instituto Nacional de Medicina Legal y Ciencias Forenses* (INMLCF) documenta los cadáveres de víctimas de muertes violentas que examina. La Policía Nacional de Colombia rastrea los homicidios y la *Fiscalía* registra las desapariciones a partir de denuncias y para efectos de juicios legales. Debemos tener presente, para efectos del presente estudio, que las víctimas que se incluyen en cada conjunto de datos varían según el mandato de cada organización.

El universo inicial de los datos que revisamos abarca un conjunto más amplio de clasificación de datos, que incluye todos los registros en Casanare clasificados como desapariciones, secuestro simple, secuestro extorsivo, secuestrado liberado, muerte, rehén, o detención. Excluimos todos los suicidios y las muertes por accidentes. A continuación cruzamos los datos como se describe en la Sección 3.2: es decir, revisamos cada registro de cada base de datos y verificamos en todas las otras bases si la misma persona figuraba varias de ellas. Tras el proceso de pareo, los registros clasificados como detenciones, rehenes y secuestros extorsivos se conservaron sólo si se podían parear con otro registro clasificado como muerte o desaparición. Todos los demás registros clasificados como detenciones, rehenes y secuestros se eliminaron del estudio.

En la Tabla 3 la columna de registros totales indica cuántos registros de homicidios y desapariciones figuraban en cada conjunto de datos; la siguiente columna delimita esos registros a los que corresponden

<sup>8</sup>Deseamos agradecer a todas las organizaciones que generosamente contribuyeron sus datos con nosotros para este informe.



al período del estudio, 1998-2007; la columna final, Únicos, indica cuántos registros de cada conjunto de datos en el período del estudio figuran sólo en esa lista y en ningún otro conjunto de datos. Por ejemplo, en el conjunto de datos de muertes del INMLCF aparece un total de 2.168 registros, de los que 2.085 corresponden al período 1998-2007, y hay 1.420 víctimas de ese mismo período que no figuran en ninguna otra lista.

Cuadro 3: Total de registros y registro únicos por Muerte y Desaparición

Nombre de la Organización	Total de registros	1998-2007	Únicos
INMLCF - Muertes violentas	2168	2085	1420
Fiscalía General de la Nación - Desaparecidos	1313	659	387
Observatorio de Derechos Humanos de la Vicepresidencia	528	501	284
Policía Nacional (DIJIN)	825	825	221
CINEP	338	274	97
Fondelibertad	332	312	67
Comisión Colombiana de Juristas	250	217	51
Coordinación Colombia Europa Estados Unidos	77	72	30
INMLCF - Desaparecidos	193	172	10
Fiscalía Seccional Santa Rosa de Viterbo - Desaparecidos	169	163	4
Organizaciones de Familiares de Víctimas	52	50	1
Gaula	128	111	1
Cuerpo Técnico de Investigaciones de la Fiscalía	36	36	0
Equitas	28	23	0
País Libre	9	9	0

La Tabla 3 deja en claro que cada conjunto de datos individual es incompleto. Sin embargo, muchos de los conjuntos de datos aportan información nueva: la mayoría de los conjuntos de datos en la Tabla 3 contienen al menos algunos hechos documentados únicamente en ese conjunto de datos. Ver en la Sección 3.4 comparaciones descriptivas de los conjuntos de datos.

### 3.2. Vinculación de registros

La vinculación de registros (record linkage en inglés) - o pareo - es la identificación de múltiples registros en una base de datos que se refieren al mismo hecho o individuo. Para efectos de las estimaciones estadísticas hay dos tipos de pareo de registros de víctimas: al interior de un sistema y entre sistemas.

Una forma de entender el pareo al interior de un mismo sistema es la eliminación de los registros duplicados. Ésta se realiza cuando la misma víctima aparece registrada más de una vez en un mismo conjunto de datos. La duplicación de registros es bastante común; varios miembros de una misma familia pueden denunciar un crimen a una misma institución, o el mismo hecho puede registrarse en múltiples fuentes de prensa. El pareo entre sistemas es la vinculación de registros a través de los diferentes conjuntos de datos. Este es el caso si un testigo denuncia un crimen a dos o más instituciones - a una ONG y a la Fiscalía, por ejemplo.

Realizamos la vinculación de registros al interior de un mismo sistema y entre sistemas simultáneamente. Cada conjunto de datos disponible para este estudio es un listado con los nombres, el sexo, la fecha, la ubicación y el tipo de violencia. Algunos de los conjuntos de datos incluyen también información sobre el supuesto perpetrador, las características físicas de las víctimas y otras variables. Esto significa que los conjuntos de datos contienen diferentes variables. Antes de vincular los registros, seleccionamos un conjunto mínimo de variables comunes para todos los conjuntos de datos. El próximo paso fue combinar todos los conjuntos de datos en uno sólo, más grande, manteniendo un identificador único para cada registro de cada una de las

bases de datos.

El Programa de Derechos Humanos de Benetech utilizó varios criterios de selección para agrupar en un conjunto los múltiples registros sobre la misma víctima. Los registros que creemos se refieren a la misma persona a veces coincidían exactamente. Otras veces coincidían en por lo menos el nombre de pila de la víctima, su primer y su segundo apellidos. Las fechas y lugares a veces eran exactos o variaban por unos días. A veces los dígitos de los meses y los días parecían estar invertidos, pero todo el resto de la información coincidía.

Para establecer la sensibilidad de la deduplicación y del pareo de registros entre sistemas, los datos fueron pareados por dos personas. Luego evaluamos la similitud en la manera como las dos agruparon los registros. Las dos comenzaron con el mismo conjunto de pares identificados automáticamente, los que modificaron identificando nuevos pares y desvinculando registros agrupados erróneamente.

Cuantificamos los resultados del pareo contando el número de pares de registros en los cuales las dos personas coincidieron y en los que no coincidieron. Por ejemplo, digamos que se encontraron registros sobre una persona con el nombre de Juan García en cinco conjuntos de datos diferentes. Después de examinar los detalles de los registros en cuanto a fechas y lugares, la primera persona realizando el pareo agrupó los cinco registros mientras que la segunda persona sólo agrupó cuatro de ellos y opinó que el quinto registro se refería a un desaparecido diferente. En seis *pares* de registros hubo acuerdo entre las dos personas. Calculamos la tasa de acuerdo entre las dos personas; este es el número de veces en que coincidieron sobre la correspondencia de dos registros de una hecho letal dividido por el número de veces en que hubieran podido coincidir. Usando este cálculo, encontramos que las dos personas coincidieron 91.1 % del tiempo. Hubo 280 pares de registros pareados por la primera pero no por la segunda persona, 149 pareados por la segunda pero no por la primera persona, y 4.389 pares de registros pareados por ambas personas.

La elevada tasa de concordancia entre las dos personas que realizaron el pareo sugiere que el resultado del pareo no depende de manera decisiva de quién lo ejecute. Esta confiabilidad en el pareo sugiere también que existen suficientes detalles en los datos para conectar los múltiples reportes sobre los mismos hechos letales de forma confiable y determinar así las tasas de traslape entre las fuentes de datos.

### 3.3. Contradicciones en el proceso de pareo

A veces los registros que claramente se referían a las mismas víctimas contenían información contradictoria. Por ejemplo, una persona puede ser clasificada como “desaparecida” en un conjunto de datos y “muerta” en un otro conjunto de datos. De la misma manera, la fecha o lugar de la desaparición puede variar debido a información reportada de manera imprecisa o a errores de digitación en el registro de los datos.

Hay tres áreas específicas en las que los registros pueden contradecirse: en materia de tipo de violencia, de municipio y de año. Creamos la siguiente jerarquía de reglas para resolver las contradicciones al interior de lo que llamamos grupos de pares - es decir, grupos de registros que emanan de diferentes conjuntos de datos y que creemos se refieren a la misma persona:

Cuando surgían discrepancias en los registros al interior de registros que se referían a la misma persona en cuanto a tipo de violencia:

- Si por lo menos un registro fue clasificado como muerte, el hecho se consideró una muerte
- Si por lo menos un registro fue clasificado como desaparición y no había una muerte en el grupo de pares, el hecho se consideró una desaparición
- Los registros clasificados como detenciones, rehenes y secuestros extorsivos se mantuvieron sólo si se podían parear con otro registro clasificado como muerte o desaparición. Todos los demás registros

clasificados como detenciones, rehenes y secuestros no se consideraron “letales” y por tal razón se excluyeron del estudio

Cuando en los registros al interior de un grupo de pares había discrepancias en cuanto al municipio:

- Si a un registro le faltaba el municipio y en el otro figuraba un municipio válido, escogimos el municipio válido;
- Si se registraban dos o más municipios, escogimos el que figuraba con más frecuencia;
- Si en un registro figuraba la capital, Yopal, como municipio y el otro tenía un municipio diferente, escogimos el municipio que no fuera Yopal;<sup>9</sup>
- Si dos registros tenían diferentes municipios de los cuales ninguno era Yopal, escogimos el municipio aleatoriamente.

Cuando hubo discrepancias en los registros dentro de un grupo de pares con respecto al año:

- Si se habían reportado dos o más años, escogimos el que se reportó con mayor frecuencia.
- De los años más frecuentes, escogimos el más reciente.

Estas reglas se aplicaron antes de calcular las estimaciones.

### 3.4. Análisis descriptivo

Las estadísticas descriptivas describen y resumen las variables en una colección de datos. Son útiles para comprender los datos observados, pero no se pueden usar para sacar conclusiones inferenciales sobre la población total.

En esta sección presentamos las estadísticas descriptivas usando los datos combinados de 15 conjuntos de datos después de ser desduplicados y pareados, y una vez que se habían resuelto las contradicciones entre los reportes múltiples.

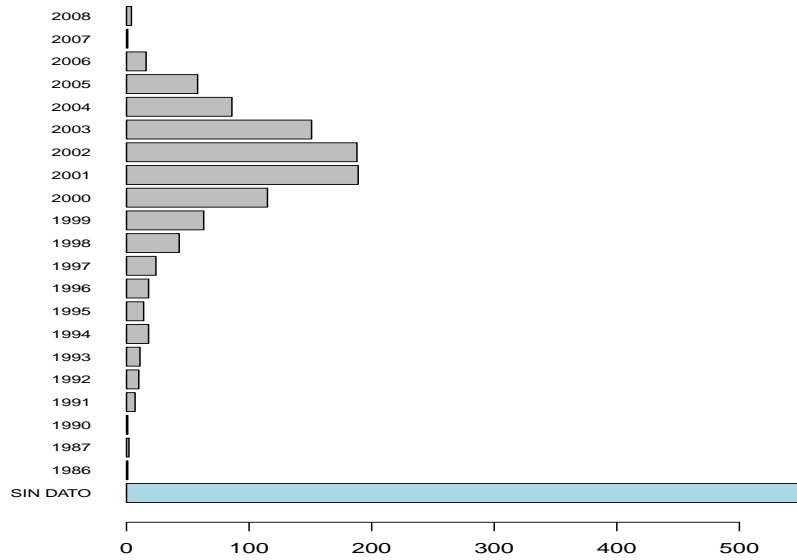
#### 3.4.1. Conteos de víctimas por tiempo y espacio

Las gráficas (abajo) del tipo de hecho por año son iguales a la barra inferior de las estimaciones por año en las Figuras 8 y 9. Las presentamos aquí para resaltar que es este el patrón de víctimas *reportadas* y para distinguir los reportes de la realidad. Como vemos en las Figuras 8 and 9, sin las estimaciones no sólo tendríamos una comprensión imperfecta de la magnitud de las muertes y las desapariciones sino que hubiéramos pasado por alto el probable aumento en los esfuerzos de los perpetradores por desaparecer a las víctimas en 2001.

---

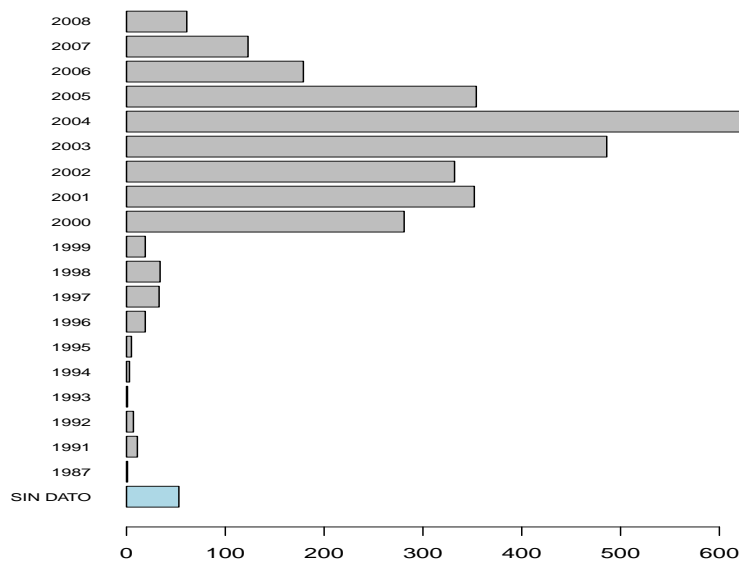
<sup>9</sup>Suponemos que, siendo probable que la gente denuncie el hecho en la ciudad capital, Yopal, donde la mayoría de las organizaciones tienen una oficina, es posible que una organización haya registrado a Yopal, el lugar de la denuncia, y no el lugar de la desaparición.

Figura 8: Conteos de desapariciones registradas en todos los conjuntos de datos por año



La Figura 8 muestra casi el mismo conteo de desapariciones en 2001-2002, pero, como vimos anteriormente, estimamos un número mayor de desapariciones no registradas en 2001. Observamos también el gran número de registros sin información sobre la fecha. Las desapariciones sin fecha de desaparición deben excluirse de la estimación porque no podemos establecer si ocurrieron durante el período del estudio

Figura 9: Conteos de homicidios registrados en todos los conjuntos de datos por año



La Figura 9 muestra un patrón de homicidios similar al patrón estimado. Llama la atención que, al contrario de lo que ocurre con las desapariciones, a relativamente pocos homicidios les falta la fecha del hecho.

A continuación presentamos los conteos de víctimas registrados por municipio. Es útil ver los datos subyacentes que se utilizaron para las estimaciones por región geográfica, especialmente dado que no se pudo calcular estimaciones por municipio. Vemos en las Figuras 10 y 11 que Yopal tiene las cifras más altas tanto de homicidios como de desapariciones, pero hay muchos registros sin información sobre el lugar, especialmente de las desapariciones.

Figura 10: Conteos de desapariciones registradas en todos los conjuntos de datos por municipio

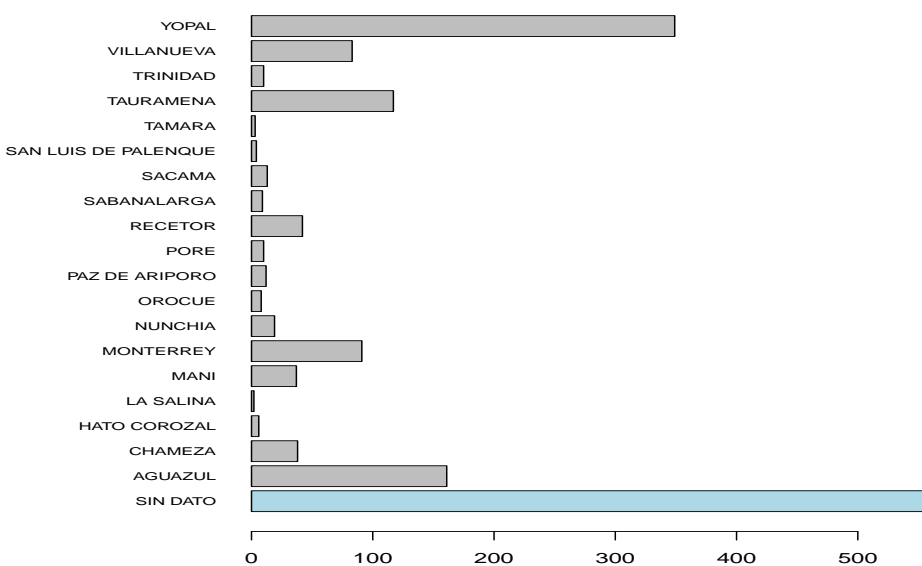
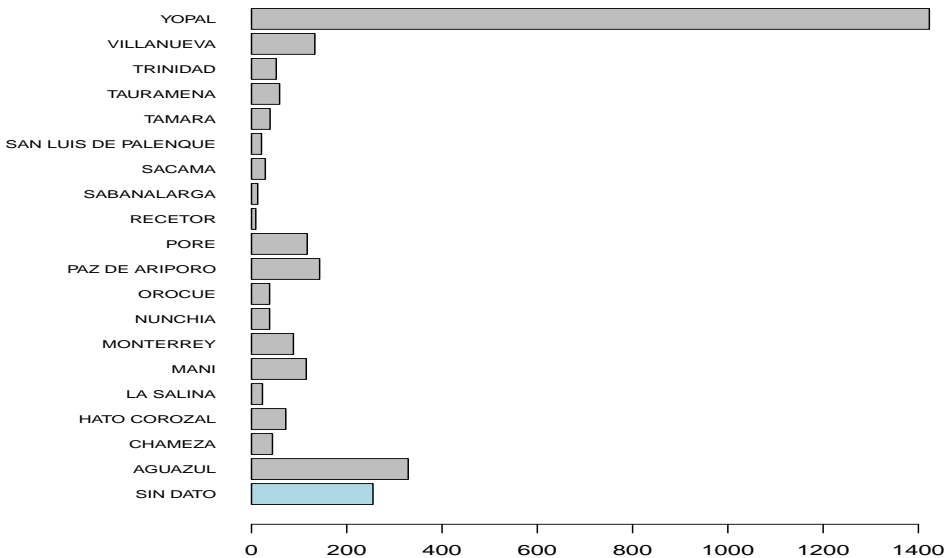


Figura 11: Conteos de homicidios registrados en todos los conjuntos de datos por municipio



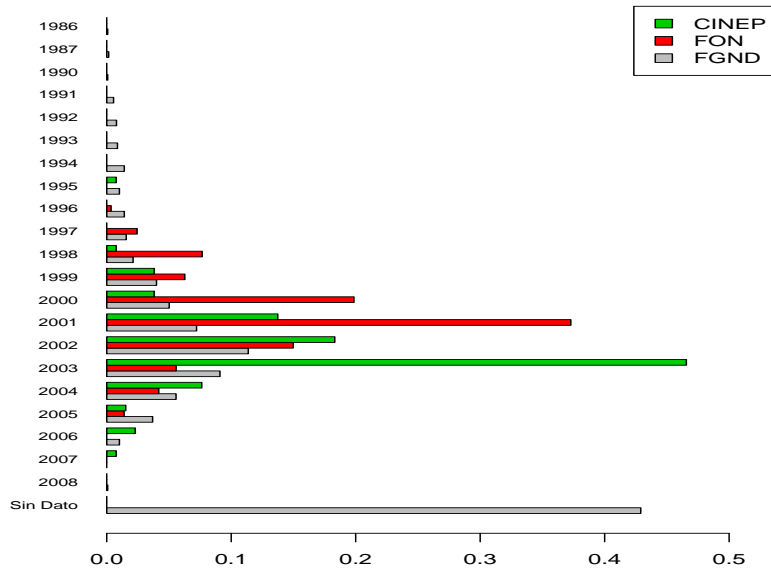
### 3.4.2. Proporción de víctimas en cada conjunto de datos por tiempo y espacio

A veces se utiliza un conjunto de datos para intentar explicar “la historia” de “lo que ocurrió”. Abajo presentamos una serie de gráficas sobre cómo esas historias pueden discrepar sutilmente - o evidentemente - dependiendo de cuál conjunto de datos se use. En los debates, los recopiladores de conjuntos de datos tienden a explicar las diferencias aduciendo a las definiciones o a manipulaciones de los datos por motivos políticos. Creemos que cada lista ha sido recopilada de la mejor manera que puede hacerlo cada organización, y que las diferencias se deben a diferentes niveles de acceso, seguridad, relaciones de confianza con la población de víctimas y recursos, incluyendo personal, logística y capacidad de procesar la información.

Las Figuras 12 - 17 presentan tres conjuntos de datos por gráfica (seleccionados por ser los conjuntos de datos más grandes que tenemos sobre ese tipo de hecho) y muestran qué proporción de cada lista incluye registros para cada año o cada municipio. En otras palabras, si estudiáramos cada lista por separado, ¿qué nos diría sobre el año o el municipio “más violento”, el “menos violento” y el patrón entre los dos? Desde luego, la historia no tiene que ver necesariamente con el verdadero patrón de violencia sino, más bien, con el patrón de observaciones de cada organización.

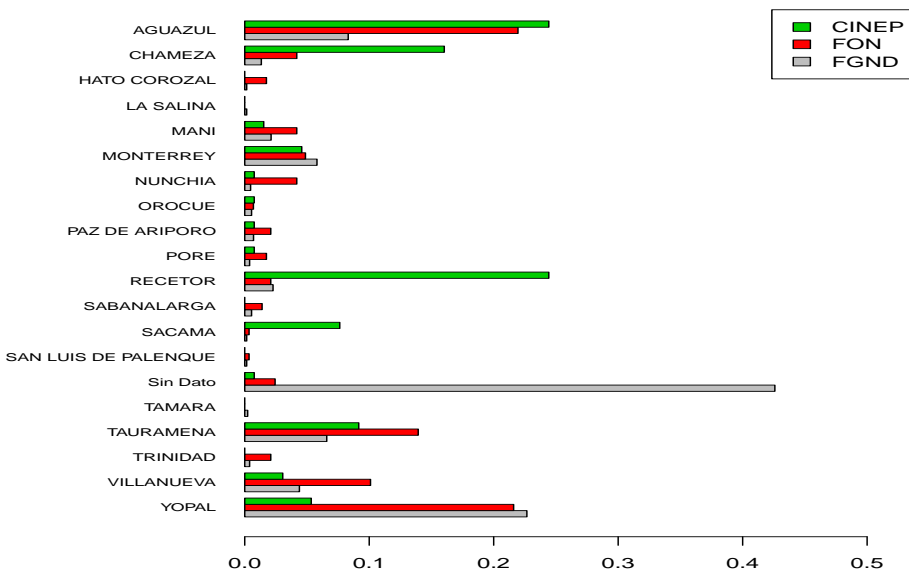
En la Figura 12 vemos que las tres listas reportan un aumento gradual por año y después una caída en años posteriores. Sin embargo, hay una gran varianza por lista por año. El pico observado por Fondelibertad (FON) ocurre en 2001 y el pico en CINEP en 2003. La mayor proporción de desapariciones registradas por la Fiscalía (FGND) carece de fechas.

Figura 12: Proporción de desapariciones por conjunto de datos por año



Al comparar los mismos tres conjuntos de datos por municipio, vemos diferencias muy interesantes en los patrones. El CINEP observa la proporción más alta de desapariciones en Aguazul y Recetor. Los otros dos muestran más violaciones proporcionalmente en Yopal y, en menor grado, en Aguazul. Una vez más, FGND muestra una muy alta proporción de datos sin información sobre el municipio. FON capta datos que los otros dos no han captado en Hato Corozal, Nunchia y Trinidad.

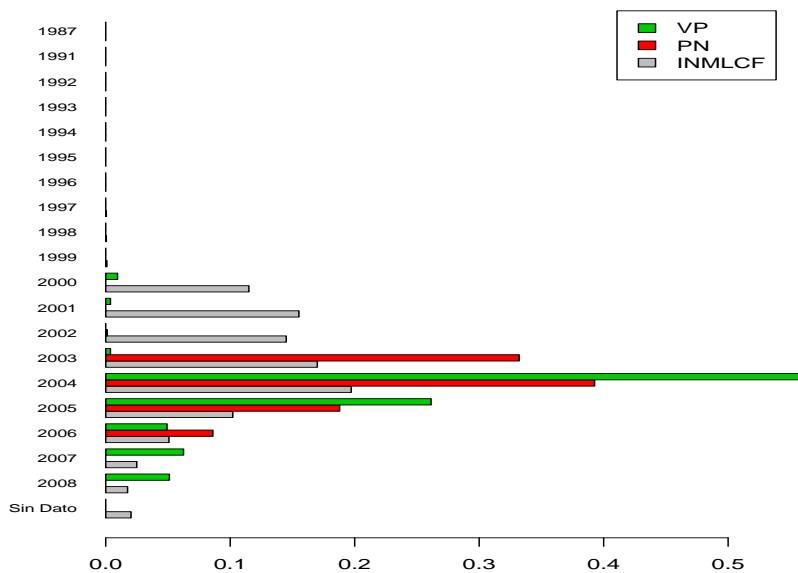
Figura 13: Proporción de desapariciones por conjunto de datos por municipio



En materia de homicidios, presentamos dos juegos de gráficas. Las Figuras 14 - 15 trazan la proporción de homicidios de los mayores conjuntos de datos, e incluyen los homicidios en general. Las Figuras 16 - 17 comparan los conjuntos de datos que registran sólo los homicidios sociopolíticos.

La Figura 14 explicita que los conjuntos de datos compartidos con nosotros no cubren los mismos años.

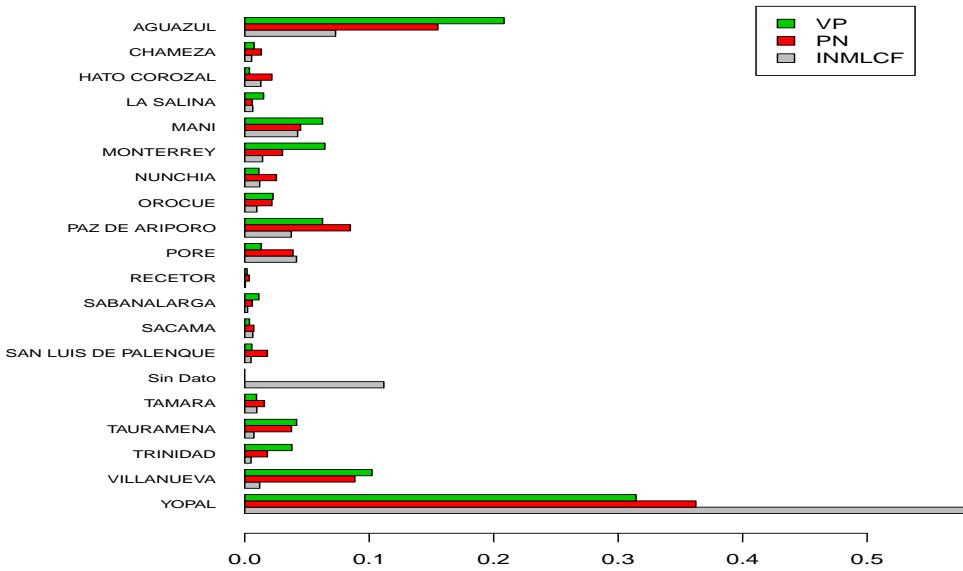
Figura 14: Proporción de homicidios por conjunto de datos por año



La Figura 15 demuestra que el INMLCF, la PN y la VP harían un recuento similar por municipio, aunque el del INMLCF sería más dramático en el caso de Yopal que los otros dos.

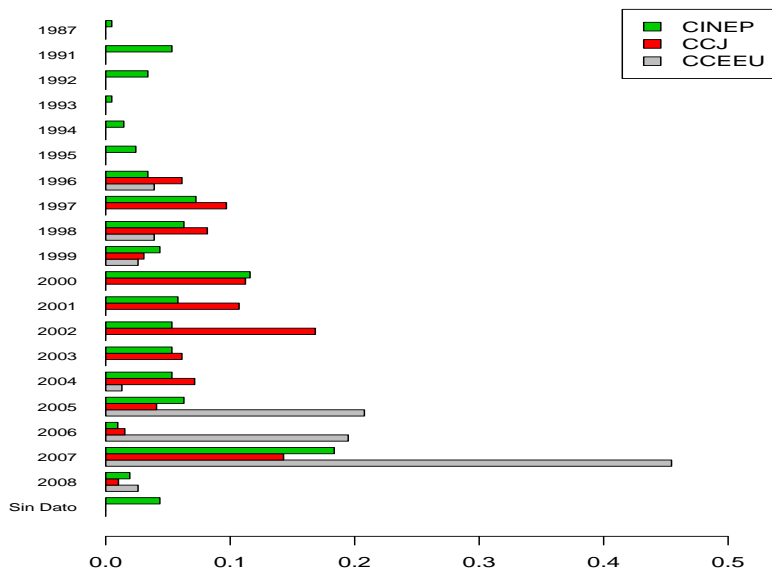


Figura 15: Proporción de homicidios por conjunto de datos por municipio



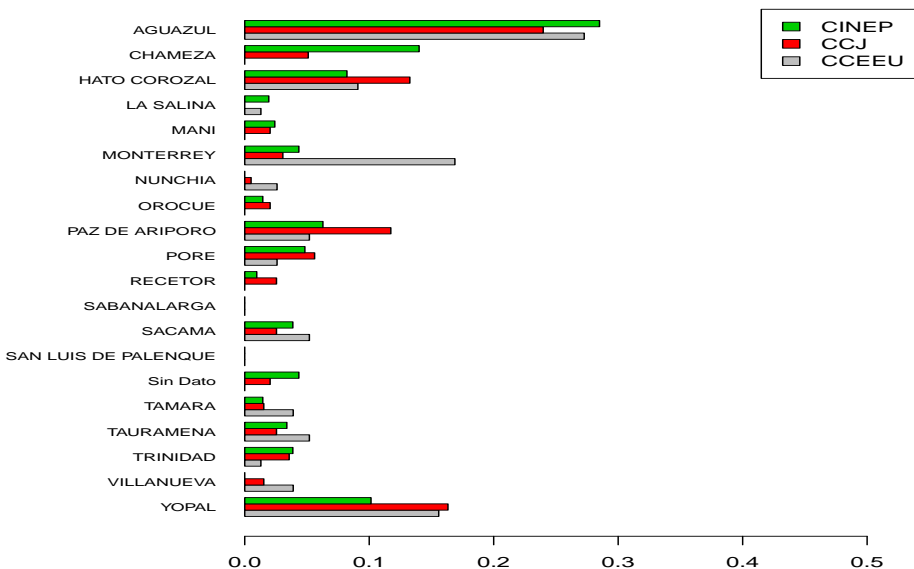
En la Figura 16, vemos que el CINEP y la Coordinación Colombia Europa Estados Unidos (CCEEU) tienen la mayor cobertura de homicidios en 2007. La Comisión Colombiana de Juristas (CCJ) exhibe proporciones similares en 2002 y 2007.

Figura 16: Proporción de homicidios políticos por conjunto de datos por año



Finalmente, en la Figura 17 vemos que las tres listas muestran cerca de 25% de sus reportes en Aguazul. La CCEEU registra una proporción relativamente alta de cobertura en Monterrey. La CCJ observa una mayor proporción en Hato Corozal y Paz de Ariporo que las otras dos listas.

Figura 17: Proporción de homicidios políticos por base de datos por municipio



Creemos que estas diferencias en reportaje eran de esperarse. Las diferencias se deben a que las diversas organizaciones se enfrentan a retos diferentes en la recolección de datos y reportaje bajo diferentes mandatos, con un grado desigual de acceso a las redes sociales y diferentes niveles de recursos. Ninguna de las listas tiene “la verdad revelada”. Todas ellas cuentan parte de la historia y todas complementan la creación del conocimiento en torno a la violencia del pasado. El análisis descriptivo es un paso analítico intermedio.

Para superar las diferencias entre los conjuntos de datos individuales, son de vital importancia los métodos estadísticos que hacen deducciones estadísticas a partir de múltiples fuentes de datos. Las fuentes de datos individuales son incompletas, pero juntas iluminan nuestra comprensión del universo de hechos violentos que se estudia. Basándose en múltiples fuentes de datos, la Estimación por Sistemas Múltiples (ESM) aporta una manera de estimar el subregistro - es decir, los hechos o individuos que pueden no aparecer registrados en ninguna fuente - y en últimas calcula el número total de hechos violentos, tanto documentados como no documentados.

## 4. Conclusión

Dado el grado de incertidumbre que aún persiste tras combinar los datos de 15 conjuntos de datos, y al calcular las estimaciones de las víctimas desconocidas y totales, desaconsejamos enfáticamente el uso de un solo conjunto de datos para hacer afirmaciones sobre las tendencias o los patrones de violencia. Las afirmaciones sobre los patrones de violencia son importantes porque a menudo apuntan a influir en las políticas, a orientar los esfuerzos de intervención y a esclarecer la historia.

Creemos que las estimaciones de la violencia letal importan históricamente por varias razones. Primero, los hechos letales no reportados representan la desaparición del conocimiento social. Las víctimas no documentadas pueden quedar completamente invisibles porque no queda una huella duradera de su muerte o desaparición. Algunas víctimas pueden permanecer invisibles aunque se recuperen sus restos si esos restos no se identifican; estos son los cadáveres N.N.. En algunos casos, la víctima puede ser la única en saber que ocurrió el homicidio o la desaparición, pues su familia puede optar por no denunciar el crimen. Nuestras estimaciones del número de homicidios y desapariciones sin contar aseguran que estas víctimas invisibles no queden excluidas de la comprensión histórica de la violencia.

## 5. Trabajo futuro

Aunque nuestro análisis de los homicidios y las desapariciones en Casanare representa un importante avance en nuestra metodología, aún se puede hacer mucho más. Estamos desarrollando más los detalles de este método en la literatura estadística académica. Nuestra intención en el futuro es aplicar esta y otras metodologías en otros departamentos y eventualmente a través de toda Colombia.

## Agradecimientos

Este proyecto fue financiado por la Agencia Sueca de Cooperación Internacional para el Desarrollo (Asdi) en Colombia y a través del apoyo institucional a Benetech de las Fundaciones John D. y Catherine T. MacArthur, Oak y Sigrid Rausing. Estamos muy agradecidos a estas entidades por su continuo apoyo.

Tuvimos la suerte de trabajar conjuntamente con EQUITAS en los aspectos de financiamiento, recolección de información y diseminación de este trabajo. Deseamos agradecer también a las quince organizaciones que pusieron sus datos sobre los homicidios y desapariciones a disposición nuestra y de nuestro socio EQUITAS. En especial, nuestro reconocimiento va a los expertos locales, quienes en diversas ocasiones compartieron con nosotros sus conocimientos cualitativos sobre el conflicto en Casanare y nos ayudaron a desarrollar estratos significativos.

Un agradecimiento especial para David Banks de Duke University y Bala Rajaratnam de Stanford University por evaluar el desarrollo metodológico de ESM que se utilizó en el presente estudio.

Finalmente, nuestros agradecimientos a los miembros del equipo del Programa de Derechos Humanos de Benetech que contribuyeron al proyecto, en particular a Jeff Klingner, por los cálculos de confiabilidad entre codificadores y por la información representada en los mapas; a Ann Harrison por su cuidadosa edición; y a Beatriz Vejarano por su traducción del informe al español y por su apoyo en muchas de las tareas cotidianas necesarias para llevar a cabo el proyecto.

## Sobre los autores

Tamy Guberek, B.A., es Coordinadora para América Latina del Programa de Derechos Humanos de Benetech. La Sra. Guberek trabaja con las copartes de Benetech en Colombia y Guatemala en la realización de análisis de datos defendibles y en la mejora de sus sistemas de manejo de información sobre derechos humanos. Tiene un grado universitario en Relaciones Internacionales y Estudios de Paz y Justicia de la Universidad de Tufts, en Estados Unidos.

Daniel Guzmán, B. S. es consultor estadístico del Programa de Derechos Humanos de Benetech. Ha contribuido al diseño de proyectos, a la capacitación de copartes y al análisis de datos en Colombia y Guatemala. También apoyó procesos de análisis de datos en proyectos en Sierra Leona y Liberia. Estudió Estadística en la Universidad Nacional de Colombia.

Megan Price, Ph.D., es estadística en el Programa de Derechos Humanos de Benetech. Ha contribuido con análisis estadísticos en proyectos en Guatemala y Colombia. La Dra. Price obtuvo su Ph.D. en Bioestadística en la Universidad de Emory, en Estados Unidos.

Kristian Lum, M.S., fue pasante en el Programa de Derechos Humanos de Benetech y actualmente es consultora estadística. Es candidata para el Ph.D. en la facultad de Estadística de la Universidad de Duke, en Estados Unidos.

Patrick Ball, Ph.D., es el Vicepresidente del Programa de Derechos Humanos en la Iniciativa Benetech. Desde 1991, el Dr. Ball ha diseñado sistemas de manejo de información y conducido análisis estadísticos para proyectos de información sobre derechos humanos a gran escala, utilizados por las comisiones de la verdad, organizaciones no gubernamentales, tribunales y misiones de Naciones Unidas en El Salvador, Etiopía, Guatemala, Haití, Sudáfrica, Kosovo, Sierra Leona, Perú, Timor-Leste y Chad. Actualmente el Dr. Ball participa en proyectos de Benetech en Colombia, Burma, Líbano, Guatemala y otros países alrededor del mundo.

## Sobre el Programa de Derechos Humanos de Benetech

El Programa de Derechos Humanos de Benetech tiene más de 17 años de experiencia aplicando análisis científicos rigurosos a datos sobre violaciones de derechos humanos. Nuestra práctica prolongada ha sido de utilidad para nueve comisiones de la verdad y reconciliación, diversas misiones de Naciones Unidas y organismos oficiales de derechos humanos, así como para tribunales penales internacionales. Hemos ejecutado proyectos en El Salvador, Etiopía, Guatemala, Haití, Sudáfrica, Kosovo, Sierra Leona, Sri Lanka, Timor-Leste, Colombia y Perú. También hemos brindado orientación sobre el procesamiento de datos y metodologías de análisis a numerosas organizaciones no gubernamentales en muchos países al rededor del mundo. Con nuestras copartes generamos argumentos científicamente defendibles basados en rigurosa evidencia. (<http://www.benetech.org>, <http://www.hrdag.org>).

Los materiales que contiene este estudio representan la opinión de sus autores y editores y no deben interpretarse como la opinión de la Iniciativa Benetech, de cualquiera de los proyectos constituyentes de Benetech, del Directorio de Benetech o de los donantes a Benetech.

Derechos de autor 2009 por  
Iniciativa Benetech  
480 S. California Ave., Suite 201  
Palo Alto, CA 94306-1609  
tel: +1 650-475-5440  
fax: +1 650-475-1066  
Email: [info@benetech.org](mailto:info@benetech.org)  
Web: <http://www.benetech.org>

Se otorgan ciertos derechos bajo la licencia denominada Creative Commons Attribution-NonCommercial-Share Alike License, disponible en la Red en el siguiente sitio: <http://creativecommons.org/licenses/by-nc-sa/1.0/legalcode>.

Los términos de la licencia se resumen a continuación:

Se permite copiar, distribuir y mostrar este trabajo, con las siguientes condiciones:

Reconocimiento (Attribution): Debe figurar siempre el autor original de este trabajo.

No Comercial (NonCommercial): No se puede utilizar este trabajo para fines comerciales sin permiso.

Se permite la distribución de obras que se derivan de este trabajo sólo bajo licencia.

## Apéndice A: Metodología

El método ESM ha sido perfeccionado para estimar poblaciones humanas en los censos<sup>10 11</sup>, y los autores del presente documento han utilizado ESM para estimar el total de la mortalidad relacionada con conflictos en varios casos.<sup>12 13 14 15</sup> Desafortunadamente, nuestra investigación sobre los hechos violentos en Colombia nos confrontó con diversos retos en el uso de la metodología de ESM, que ya habíamos implementado previamente. Estos retos surgieron de un exitoso esfuerzo de recolección de datos en el departamento colombiano de Casanare que produjo 15 conjuntos de datos independientes documentando hechos violentos. La siguiente sección esboza el método que ya habíamos implementado anteriormente en Casanare, y la Sección 5 abordará los detalles del método actual.

### Implementación anterior del método ESM en Casanare

La versión más básica de ESM se basa en cuatro supuestos:

1. Las muestras vienen de un *sistema cerrado*:  $N$  debe referirse a la misma población en cada conjunto de datos.
2. Las unidades son *homogéneas*: cada individuo en la población tiene la misma probabilidad de captura.
3. Los sistemas son *independientes*: la probabilidad de captura en el sistema A no influye en la probabilidad de captura en el sistema B y viceversa.
4. Existe el *pareo perfecto*: es posible dividir con precisión todos los individuos captados en categorías: aquellos captados sólo por el sistema A, sólo por el sistema B y aquellos en ambos sistemas, A y B; todos los registros que se refieren a la misma unidad deben ser reconocibles como tal.

El primer supuesto, que el objeto de medición es un sistema cerrado - ya sea la población de personas en un país o la población de hechos violentos que ocurrieron en un estado - se satisface fácilmente con nuestros datos. Este supuesto generalmente no presenta mayores problemas en los datos retrospectivos sobre eventos violentos, pues los eventos que ya ocurrieron no pueden “des-ocurrir” más tarde. En otras palabras, durante el año 2010 nadie puede morir en el año 2009: la población de muertos de 2009 está cerrada.

El segundo supuesto, la homogeneidad de la posibilidad de captura, probablemente no vale para ningún tipo de datos sobre violencia. Por ejemplo, las personas con menores contactos sociales pueden tener a la vez una mayor probabilidad de experimentar un acto violento, como un secuestro, y una menor probabilidad de ser registradas como víctimas. Una manera de abordar la probabilidad desigual de captura es por medio de la estratificación. Intuitivamente tiene sentido estratificar tanto por espacio como por tiempo, puesto que es probable que tanto el área geográfica diferente como los períodos diferentes a lo largo del tiempo tengan diferentes probabilidades de captura. Los detalles específicos de la estratificación para Casanare se encuentran en la Sección 2.1. Una estratificación eficaz requiere que en cada estrato haya suficientes datos en todos los

---

<sup>10</sup>John Darroch, Stephen Fienberg, Gary Glonek, and Brian Junker. A three-sample multiple-recapture approach to census population estimation with heterogeneous catchability. *Journal of the American Statistical Association*, 88 (423):1137-1148, 1993.

<sup>11</sup>Chandra C. Sekar and Edwards W. Deming. “On a method of estimating birth and death rates and the extent of registration”. *Journal of the American Statistical Association*, 44(245):101-115, 1949.

<sup>12</sup>Patrick Ball. “Political Killings in Kosova/Kosovo March-June 1999”. American Bar Association, 2000.

<sup>13</sup>Patrick Ball. “Making the Case: Investigating Large Scale Human Rights Violations Using Information Systems and Data Analysis”, ch. 12. *American Association for the Advancement of Science*, Washington, DC, 1999.

<sup>14</sup>Patrick Ball, Jana Asher, David Sulmont, and Daniel Manrique. “How Many Peruvians Have Died? An estimate of the total number of victims killed or disappeared in the armed internal conflict between 1980 and 2000”. Report to the Peruvian Commission for Truth and Reconciliation (CVR), Washington, DC, August 2003.

<sup>15</sup>Romesh Silva and Patrick Ball. “The Profile of Human Rights Violations in Timor-Leste, 1974-1999. Report by the Benetech Human Rights Data Analysis Group to the Commission on Reception, Truth and Reconciliation (CAVR)”, 2006.

sistemas y suficiente traslape entre sistemas.

El tercer supuesto, la independencia de los sistemas, es igualmente difícil de cumplir. Un ejemplo común de ello es la diferencia entre organizaciones gubernamentales y no gubernamentales. Dado que las diferentes poblaciones pueden sentir diferentes niveles de confianza hacia los dos tipos de organizaciones, es poco probable que un testigo que informe a un tipo de organización informe al otro tipo: la probabilidad de captura de un sistema afecta la probabilidad de captura del otro.

El cuarto supuesto, el pareo perfecto entre sistemas, conduce a la parte más intensiva en cómputos del proceso de sistemas múltiples. Actualmente no existen modelos tratables de ESM con pareo imperfecto; la tarea consiste entonces en parear los registros con la mayor precisión posible utilizando algún(os) identificador(es) único(s). Para mayores detalles sobre este proceso en Casanare, ver la Sección 3.2.

Resumiendo, se puede afirmar que los supuestos uno y cuatro se cumplen con nuestros datos. En el caso del supuesto dos, el ajuste se hace por medio de la estratificación. En el informe pasado sobre Casanare, el ajuste para el supuesto tres se hizo modelando las dependencias entre listas utilizando modelos loglineales como lo sugieren Bishop et al.<sup>16</sup> Desafortunadamente, la multitud de listas disponibles en Casanare imposibilita el uso de este método (es computacionalmente intratable enumerar cada posible modelo loglineal según el método que sugieren Bishop et al.).

Así, tuvo que desarrollarse una nueva metodología para estimar el número total de víctimas de violencia letal en Casanare. La siguiente sección esboza este método.

## La implementación actual de ESM en Casanare

Uno de los primeros retos por superar es la necesidad de reducir los datos. No es posible enumerar todos los modelos posibles según el enfoque loglineal, pero tampoco queremos descartar ningún dato. Dado que la teoría y la metodología de ESM están bien desarrolladas para tres sistemas (listas o conjuntos de datos) optamos por examinar toda combinación posible de tres listas de nuestro total de 15. El resultado es  $\binom{15}{3} = 455$  posibles divisiones. Pero en la práctica resultaron menos divisiones “válidas”. Una división se consideraba válida si los tres conjuntos de datos en esa división específica contenían suficientes traslapes para calcular una estimación utilizando ESM.

Como se mencionó en la Sección 2.2, para algunos estratos no había suficientes datos para calcular estimaciones. “Datos insuficientes” implica una de dos cosas: 1) cada conjunto de datos observó sólo un pequeño número de hechos en un estrato dado y por ende hay poco o ningún traslape entre conjuntos de datos; o 2) los conjuntos de datos registraron muchos hechos diferentes y por ende no queda ningún traslape. La segunda perspectiva, muchos eventos diferentes, indica la posibilidad de que la población no observada de hechos letales sea, en realidad, considerable. No obstante, utilizando los datos disponibles no hay manera de generar una estimación plausible para estos casos de traslape insuficiente. Por lo tanto, al omitir estos casos podríamos estar subestimando el número total de hechos letales en Casanare - i.e. nuestras estimaciones podrían ser conservadoras. Creemos que nuestras estimaciones tienen un sesgo hacia abajo en aquellos estratos para los cuales los datos disponibles son muy limitados, o para los cuales los traslapes observables son limitados.

Al interior de cada división podemos construir también varios modelos posibles para dar razón de las posibles dependencias entre las listas. El siguiente reto es decidir qué modelo utilizar. A menudo se utiliza el Criterio Bayesiano de Información (BIC en inglés) para indicar la bondad de ajuste de un modelo, y se

---

<sup>16</sup>Yvonne M. M. Bishop, Stephen Fienberg, and Paul H. Holland. *Discrete Multivariate Analysis: Theory and Practice*. MIT Press, Cambridge, MA, 1975.

selecciona el modelo con el menor BIC como el ‘mejor’ modelo.

No obstante, en muchos casos existen muchos modelos plausibles que producen estimaciones muy diferentes. El método Bayesian Model Averaging (BMA) toma en consideración la incertidumbre sobre el modelo mismo. El supuesto bayesiano subyacente es que cada uno de los parámetros del modelo y el modelo como tal son variables aleatorias y no valores fijos. Para conocer los parámetros y los modelos, hay que saber algo sobre su distribución. Es decir, un bayesiano quisiera saber sobre  $P(N|\mathbf{Y})$ , la probabilidad de cualquier  $N$  (el número total estimado de hechos observados y no observados) dados nuestros datos observados  $\mathbf{Y}$ . Cada elemento de nuestros datos observados  $\mathbf{Y}$  es la suma del número total de nombres que figuran en cada intersección de las listas.  $P(N|\mathbf{Y})$  puede calcularse como sigue:

$$P(N|\mathbf{Y}) = \sum_{\mathcal{M}} p(N|\mathbf{Y}, M)p(M|\mathbf{Y}), \quad (1)$$

donde  $P(N|\mathbf{Y}, M)$  es la distribución de los parámetros del modelo desconocido, dados los datos observados y cualquier modelo candidato  $M$  del conjunto de todos los posibles modelos  $\mathcal{M}$  y  $p(M|\mathbf{Y})$  es la probabilidad posterior del modelo  $M$ .

Una pieza clave de esta ecuación,  $p(M|\mathbf{Y})$  puede reescribirse como:

$$p(M|\mathbf{Y}) = \frac{p(\mathbf{Y}|M)p(M)}{\sum_{\mathcal{M}} p(\mathbf{Y}|M)p(M)} \quad (2)$$

donde  $p(M)$  es la probabilidad *a priori* del modelo de  $M$  y  $p(\mathbf{Y}|M)$  es la probabilidad marginal de los datos. Desafortunadamente,  $p(\mathbf{Y}|M)$  sólo tiene una solución analítica en circunstancias especiales. Por lo tanto, nosotros definimos el modelo para los datos en una manera que es analíticamente útil.

Usamos una distribución de probabilidad multinomial para los datos y una distribución *a priori* Dirichlet para los parámetros del modelo. Debido a una propiedad llamada conjugación entre la probabilidad multinomial y los *a priori* de Dirichlet, podemos derivar la solución analítica para  $p(\mathbf{Y}|M)$ .

$$\begin{aligned} Y|\theta, N, M &\sim \text{Multinomial}(\theta) \\ \theta|M &\sim \text{Dirichlet} \\ Y|N, M &\propto \int_{\theta} [Y|\theta, N, M][\theta|M] \end{aligned}$$

donde  $\theta$  es un conjunto de probabilidades para cada conteo en  $Y$ .

También podemos asignar una distribución *a priori* al número total de hechos letales en Casanare  $N$ , observados mas observados. El valor elegido se llama valor de truncamiento *a priori*; usamos este valor *a priori* no solamente para la conveniencia del cálculo, sino también porque creemos que el número total de hechos letales en Casanare no puede ser mayor que una proporción razonable de toda la población de Casanare. Por lo tanto, como se sugiere en Madigan y York<sup>17</sup> usamos el *a priori*,  $\pi(N) \propto N^{-1}1[N_* \leq N \leq N^*]$  que tiene probabilidad positiva sólo en  $[N_*, N^*]$  con  $N_*$ , el total de personas únicas que aparecen en el conjunto de datos (porque sabemos que por lo menos hay tantas víctimas como hechos reportados) y  $N^*$ , un límite superior para el número total de posibles hechos letales en la región, determinado por expertos en el campo.

Ahora que tenemos un modelo totalmente especificado, podemos empezar a hablar de la distribución *a posteriori* de  $N$ . Para un modelo dado, si quisiéramos conocer la probabilidad *a posterior* de cualquier  $N$ , sólo tenemos que calcular:

---

<sup>17</sup>Madigan, David and Jeremy C. York. 1997. “Bayesian Methods for Estimation of the Size of a Closed Population” *Biometrika*. 84(1), 19-31



$$P(N|\mathbf{Y}, M) = \frac{P(Y|N, M)\pi(N)}{\sum_{N=N_*}^{N^*} P(Y|N, M)\pi(N)}$$

Sin embargo, como se describió anteriormente, tenemos que promediar a través de todos modelos posibles.

$$P(N|\mathbf{Y}) = \frac{\sum_{M \in \mathcal{M}} P(Y|N, M)\pi(N)\pi(M)}{\sum_{N=N_*}^{N^*} \sum_{M \in \mathcal{M}} P(Y|N, M)\pi(N)\pi(M)}$$

Simplemente establecemos  $\pi(M) \sim |\mathcal{M}|^{-1}$ , donde  $\mathcal{M}$  es el conjunto de todos los modelos posibles.

Un último paso de complejidad, como se mencionó anteriormente, no tenemos sólo tres conjuntos de datos; tenemos 15. Sea  $\mathbf{Y}^\gamma$  la  $\gamma$ th combinación de tres conjuntos de datos y  $\mathcal{M}(\gamma)$  la representación de todos los del modelo posibles por cada combinación de tres conjuntos de datos indexadas por  $\gamma$ . Ahora, promediando sobre cada una de estas combinaciones, todo lo que queda es volver a unir las piezas como antes:

$$P(N|Y) = Z^{-1} \sum_{\gamma} \sum_{M \in \mathcal{M}(\gamma)} P(Y^\gamma|N, M)\pi(N)\pi(M, \gamma)$$

$$Z = \sum_{N=N_*}^{N^*} \sum_{\gamma} \sum_{M \in \mathcal{M}(\gamma)} P(Y^\gamma|N, M)\pi(N)\pi(M, \gamma)$$

Los resultados aplicando este metodo a los datos disponibles a nosotros en Casanare se presentan en la Sección 2.

Una de las principales diferencias en términos de interpretación de los resultados de este método, en comparación con nuestro método anterior en el informe de 2007, es la forma de medir la incertidumbre. Métodos totalmente Bayesianos, tales como los descritos en esta sección, tienen asociados intervalos de *credibilidad* del 95%. Estos son diferentes a los intervalos de *confianza*. En concreto, los intervalos de *credibilidad* se basan en una distribución totalmente especificada, lo que significa que es 95% seguro que el número correcto de los actos de violencia se encuentra dentro del intervalo (dado que el intervalo se basa en los datos observados). Esto es mucho más directo que la interpretación de la incertidumbre con respecto a los intervalos de *confianza* tradicionales. Intervalos de *confianza* tradicionales nos dicen que si llevamos a cabo el estudio 100 veces, es de esperar que en 95 de estos, la respuesta correcta se encuentre en el rango reportado.

## **Apéndice B: Descripción de las Fuentes**

### **El Instituto Nacional de Medicina Legal y Ciencias Forenses (INMLCF-M y INMLCF-D)**

El Instituto Nacional de Medicina Legal y Ciencias Forenses tiene presencia institucional directa en cerca de 166 ciudades y municipios del país, a través de 190 puntos de atención. Tiene ocho Regionales, 25 Seccionales y Unidades Básicas, algunas de estas últimas denominadas Unidades Móviles, que equivale a un cubrimiento del 65 % de la población. La información directa sobre hechos fatales viene de la actividad pericial desarrollada en cada punto de atención institucional y son almacenadas en la herramienta computarizada SIRDEC, (Sistema de Información de la Red de Cadáveres). De otra parte, el Instituto recibe la información estadística de los médicos adscritos al sector salud que se desempeñan como forenses en los casos necesarios; ésta información es conocida como estadística indirecta. La actividad forense produce información que es recolectada, procesada, analizada y divulgada, a través de la División de Referencia de Información Pericial (DRIP-CRNV). Una vez almacenada dicha información, se realiza el proceso de validación del dato en el nivel local, regional y nacional, por los funcionarios de la unidad básica, los Profesionales de Apoyo al DRIP-CRNV en las Regionales y los Técnicos de la División de Referencia de Información Pericial, DRIP-CRNV, respectivamente, para verificar que los registros estén completos y detectar posibles inconsistencias.

### **Fiscalía General de la Nación (FGND y FSR)**

La Fiscalía General de la Nación es una entidad de la rama judicial del poder público con plena autonomía administrativa y presupuestal, cuya función está orientada a brindar a los ciudadanos una cumplida y eficaz administración de justicia. Una de sus secciones principales es la Dirección Nacional de Fiscalías que a su vez tiene varias Unidades Nacionales, una de las cuales es la Unidad Nacional contra el Secuestro y la Extorsión. También existen Direcciones Seccionales, una de las cuales es la Fiscalía Santa Rosa de Viterbo (FSR): En Santa Rosa de Viterbo se documentan las denuncias de desaparecidos y secuestrados del departamento de Casanare y Boyacá, entre otros.

### **Observatorio del Programa Presidencial de Derechos Humanos y Derecho Internacional (VP)**

El Observatorio del Programa Presidencial de Derechos Humanos y Derecho Internacional mantiene una base de datos de homicidios de personas que pertenecen a “grupos vulnerables” en todo el país, tales como alcaldes y exalcaldes, concejales, periodistas, indígenas, docentes y sindicalistas. La información sobre grupos vulnerables proviene de múltiples fuentes, tales como la Federación Colombiana de Municipios, la Federación Nacional de Concejos (Fenacon), la Fundación para la Libertad de Prensa (Flip), organizaciones indígenas, el Ministerio de la Protección Social, la Policía Nacional y el DAS, entre otros. El Observatorio también realiza una Bitácora semanal de prensa que es el producto de una revisión diaria tanto de periódicos de circulación nacional y regional, como de las cadenas radiales de Caracol y RCN, que se encuentran publicados en la Internet. De dichos medios, se extrae la información relacionada con actividad política, actividad judicial, otros pronunciamientos, acciones de la Fuerza Pública y otros servidores públicos, acciones de los grupos armados al margen de la ley, otros eventos y hechos de paz y manifestaciones contra la violencia.

### **Policía Nacional (PN)**

La DIJIN (Dirección Central de Policía Judicial) es una dependencia de la Policía Nacional de Colombia. La DIJIN es responsable de las investigaciones judiciales y criminales para apoyar la administración de justicia. En años recientes la Policía Nacional se ha esforzado por mejorar su sistema de información incluyendo nuevas variables, entre ellas los nombres de las víctimas del homicidio. Reúne sus datos a nivel nacional a través de sus agentes y puestos en todo el país. Los homicidios registrados en la base de datos de la DIJIN

son aquellos que las comandancias regionales de policía conocen y sobre los cuales informan. Los datos de la Policía Nacional incluyen homicidios debidos a muchas causas, incluyendo las muertes en combate.

### **Centro de Investigación y Educación Popular (CINEP)**

El Banco de Datos recauda, sistematiza y difunde información sobre las violaciones más graves a los derechos humanos fundamentales así como sobre infracciones al Derecho Internacional Humanitario. Aunque el Banco de Datos está centralizado en Bogotá, cada vez más trabaja mediante la creación de unidades regionales en diversas zonas del país, con el fin de acercarse más al entorno de las víctimas, de sus familias y organizaciones y así lograr una información más directa y cualificada sin tener que limitarse a fuentes de prensa. En general, el Banco privilegia la verdad de las víctimas y la salvaguarda en relatos de memoria por encima de las cifras.

### **Fondelibertad (FON)**

Fondelibertad o Fondo Nacional para la Defensa de la Libertad Personal, del Ministerio de Defensa Nacional es la entidad encargada de administrar los recursos para combatir los delitos contra la libertad personal y el manejo de estadísticas de secuestro y extorsión a nivel nacional. Desde el año 1996, cuantifica la situación de las víctimas de secuestro. Su información proviene de las entidades oficiales pertenecientes a la lucha antisequestro (Fuerzas Militares, Policía Nacional, Fiscalía, DAS y CTI). Esta base de datos incluye información sobre secuestro simple, extorsivo y político.

### **Comisión Colombiana de Juristas (CCJ)**

El área de investigación de la CCJ mantiene una base de datos de violencia sociopolítica con el registro diario y permanente sobre las violaciones de los derechos humanos e infracciones del derecho internacional humanitario. Por “violencia sociopolítica”, la CCJ se refiere a actos que constituyen atentados contra la vida, la integridad personal y la libertad personal causados por los abusos de poder de los agentes del Estado; aquellos motivados por razones políticas; aquellos que se derivan de la discriminación contra las personas socialmente marginales; o aquellos causados por el conflicto armado interno. En el momento, la información procesada por la CCJ se origina en las siguientes fuentes: veinte diarios nacionales y regionales; dos semanarios de noticias de circulación nacional; denuncias directas recogidas por la CCJ; denuncias dirigidas a otras organizaciones de derechos humanos, nacionales y regionales, que hacen seguimiento de la situación de derechos humanos y derecho humanitario; información que proveen las autoridades nacionales (la Defensoría del Pueblo, la Fiscalía General, la Procuraduría General, las fuerzas estatales) y la publicación Noche y Niebla del Banco de Datos de CINEP/Justicia y Paz.

### **Coordinación Colombia Europa Estados Unidos (CCEEU)**

La Mesa de Trabajo sobre Ejecuciones Extrajudiciales de la CCEEU ha recopilado denuncias de los hechos presentadas por los familiares o las organizaciones de las víctimas. La Mesa de Trabajo recibe, sistematiza y reporta dichas denuncias como registros de ejecuciones extrajudiciales. Cada denuncia reportada siempre se contrasta con al menos un registro escrito, una denuncia pública o constancia de alguna actuación judicial o disciplinaria que la soporte.

### **La Asociación De Familiares De Detenidos-Desaparecidos- Asfaddes y Familiares Colombia (FAM)**

La Asociación De Familiares De Detenidos-Desaparecidos- Asfaddes es una organización de víctimas cuya misión desde 1982 es luchar coordinadamente con todos los asociados y organizaciones afines, por encontrar los Detenidos - Desaparecidos, exigiendo al Estado Colombiano la plena vigencia y la garantía del derecho a conocer la Verdad, la aplicación de la Justicia, la Reparación Integral y la Recuperación de la Memoria

Histórica. Por su parte Familiares de Desaparecidos Forzadamente por el Apoyo Mutuo - Familiares Colombia, fundada en el 2005, acompaña y orienta a las familias de personas desaparecidas, transmitiendo aprendizajes desde el apoyo mutuo y reforzando iniciativas de las familias en su recorrido. También adelanta acciones legales en sus casos. Ambas organizaciones llevan bases de datos con los casos que conocen y representan, o con denuncias que han recibido.

### **Los Grupos de Acción Unificada por la Libertad Personal (GAULA)**

Los GAULA son grupos élite, conformados por personal de la Policía Nacional y de las Fuerzas Militares, creados por la Ley 282 de 1996 quienes llevan a cabo operaciones de rescate de secuestrados y desmantelamiento de las bandas de criminales causantes de los delitos que menoscaban la libertad personal en Colombia. Trabajan en común acuerdo con otros organismos del Estado como el DAS y la Fiscalía. Cada uno está liderado por una Dirección Unificada compuesta por un Fiscal Especializado (secuestro y extorsión) delegado ante el GAULA y un Comandante Operativo, que pertenece a la Fuerza que constituye el componente operativo del GAULA. Ellos documentan denuncias realizadas en sus oficinas en su propia base de datos.

### **Cuerpo Técnico de Investigación de la Fiscalía (CTI)**

Otra de las secciones principales de la Fiscalía es el Cuerpo Técnico de Investigación quien planea, organiza, dirige, controla y ejecuta las funciones de Policía Judicial de la Fiscalía y asesora la definición de políticas y estrategias asociadas en los temas de investigación criminal, servicios forenses, de genética y en la administración de la información técnica y judicial que sea útil para la investigación penal. El CTI apoya las labores de búsqueda de personas desaparecidas y N.N. a través del Programa Nacional de Identificación a Víctimas N.N. y Búsqueda de Personas Desaparecidas de la Fiscalía General de la Nación. Fue creado mediante Directiva Presidencial 05 del 28 de diciembre de 1991. Este Grupo está adscrito a la Sección Nacional de Identificación de la División Criminalística y su objetivo primordial es dar nombre a todos aquellos cadáveres N.N. y lograr la ubicación de las personas desaparecidas. Este programa mantiene una base de datos con la cual busca esclarecer la identidad de personas desaparecidas.

### **EQUITAS**

EQUITAS es una organización científica y humanitaria que brinda a familias víctimas de violaciones graves, masivas o sistemáticas de los derechos humanos y en conflictos armados, contribuciones científicas e independientes para el avance de sus casos. Desde el año 2005, EQUITAS ha asesorado en temas de búsqueda, recuperación e identificación de personas desaparecidas, a un grupo de familias del departamento de Casanare que representan un total de 45 víctimas de desaparición forzada, reportadas entre 2000 y 2005, principalmente de los municipios de Monterrey, Tauramena y Villanueva. Todos estos casos se encuentran en investigación, y son representados legalmente por diferentes organizaciones de Derechos Humanos. La base de datos de estos casos tiene fines de seguimiento del proceso de búsqueda de los cuerpos de las víctimas y su restitución a sus familiares.

### **La Fundación País Libre (PL)**

La Fundación País Libre es una organización no gubernamental sin ánimo de lucro que brinda gratuitamente asistencia integral a las víctimas directas e indirectas del secuestro, la extorsión y la desaparición forzada. Recibe denuncias de secuestro simple y desaparición forzada. Sistematiza los casos atendidos en una base de datos. Participa en procesos de búsqueda y recuperación de las personas desaparecidas.