

La 'Guerra Contra las Drogas' en México: Registros (oficiales) de eventos durante el periodo de diciembre 2006 a noviembre 2011"

Laura Atuesta

Programa de Política de Drogas, Centro de Investigación y Docencia Económicas, CIDE

Oscar S. Siordia

Centro de Investigación en Geografía y Geomática "Ing. Jorge L. Tamayo", CentroGeo

Alejandro Madrazo Lajous

Programa de Política de Drogas, Centro de Investigación y Docencia Económicas, CIDE

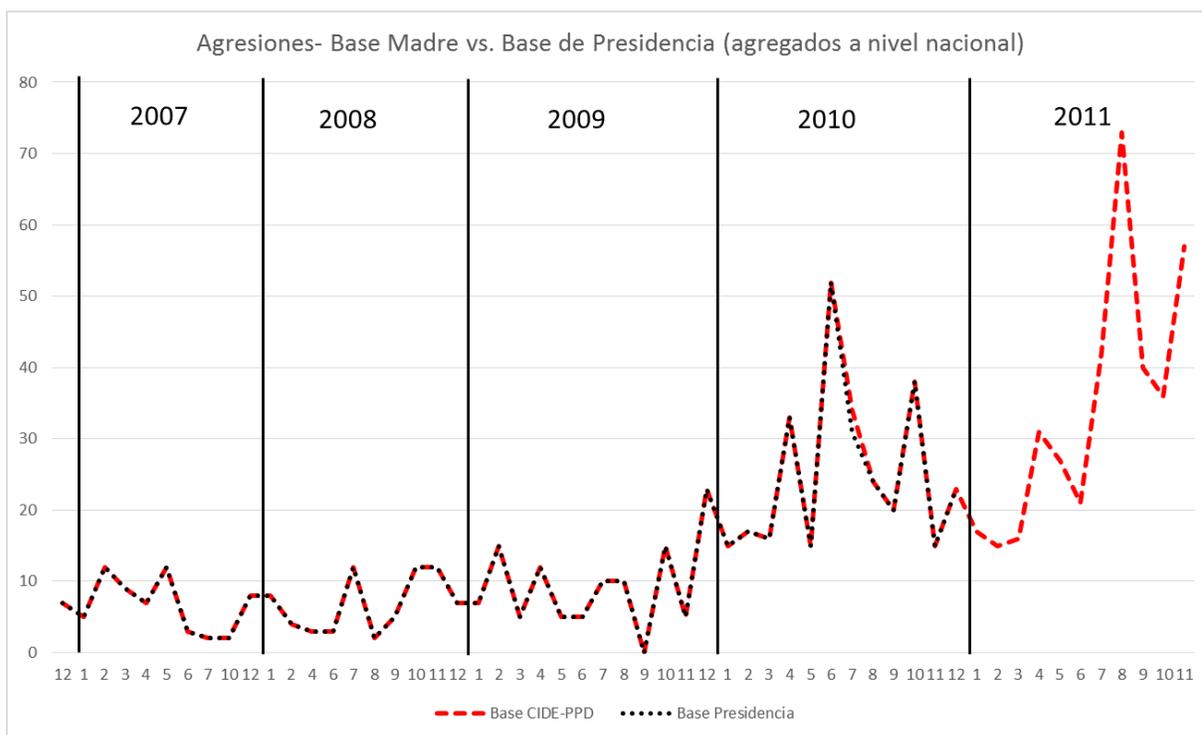
I. Introducción

A inicios del sexenio de Enrique Peña Nieto, el Programa de Política de Drogas del CIDE recibió, anónimamente, un disco que contenía diversos archivos. Al abrirlos encontramos que contenía tres bases de datos, en las que se registraban eventos violentos que tuvieron lugar en el contexto de la llamada “guerra contra las drogas” durante el gobierno de Felipe Calderón, específicamente, entre diciembre de 2006 y noviembre de 2011. De ser fidedigna, la información allí contenida permitiría comprender, a partir de hechos concretos, el desarrollo de la llamada guerra contra las drogas lanzada al inicio del sexenio anterior. Sin embargo, el anonimato de su fuente y la riqueza de su información nos obligaban a proceder con cautela. Iniciamos así un largo proceso de reflexión, discusión y tratamiento de los archivos con la finalidad de, primero, validar la información; y luego, preparar la base de datos para ser utilizada por el público en general sin comprometer la seguridad o reputación de las personas involucradas en los eventos en ella registrados. En los meses subsecuentes, un equipo multidisciplinario del Programa de Política de Drogas en colaboración con CentroGeo, diseñó un sistema de validación y codificación de la información contenida en las bases. Los resultados de los procesos de validación y codificación fueron satisfactorios y, en consecuencia, hoy hacemos pública la base de datos que llamaremos Base CIDE-PPD, a fin de que pueda ser utilizada por cualquiera que quiera investigar la política de drogas durante el sexenio de Felipe Calderón con base en los hechos que constituyen el núcleo de la llamada “guerra contra las drogas”.

Este artículo pretende describir las tres bases de datos incluidas en la Base CIDE-PPD, precisando sus limitaciones y características más importantes. Además, ofrecemos las razones que nos han llevado a concluir cuál es el origen de la Base, así como la comparación entre la información allí contenida y otros registros públicos, con el propósito de valorar la precisión de la información que se ofrece. A fin de facilitar la comprensión del análisis, llamaremos Base Madre al conjunto de las tres bases de datos que recibimos anónimamente; y Base CIDE-PPD, a la base de datos que hoy hacemos pública, resultado del ejercicio sistemático de validación y codificación de la información original contenida en la Base Madre.

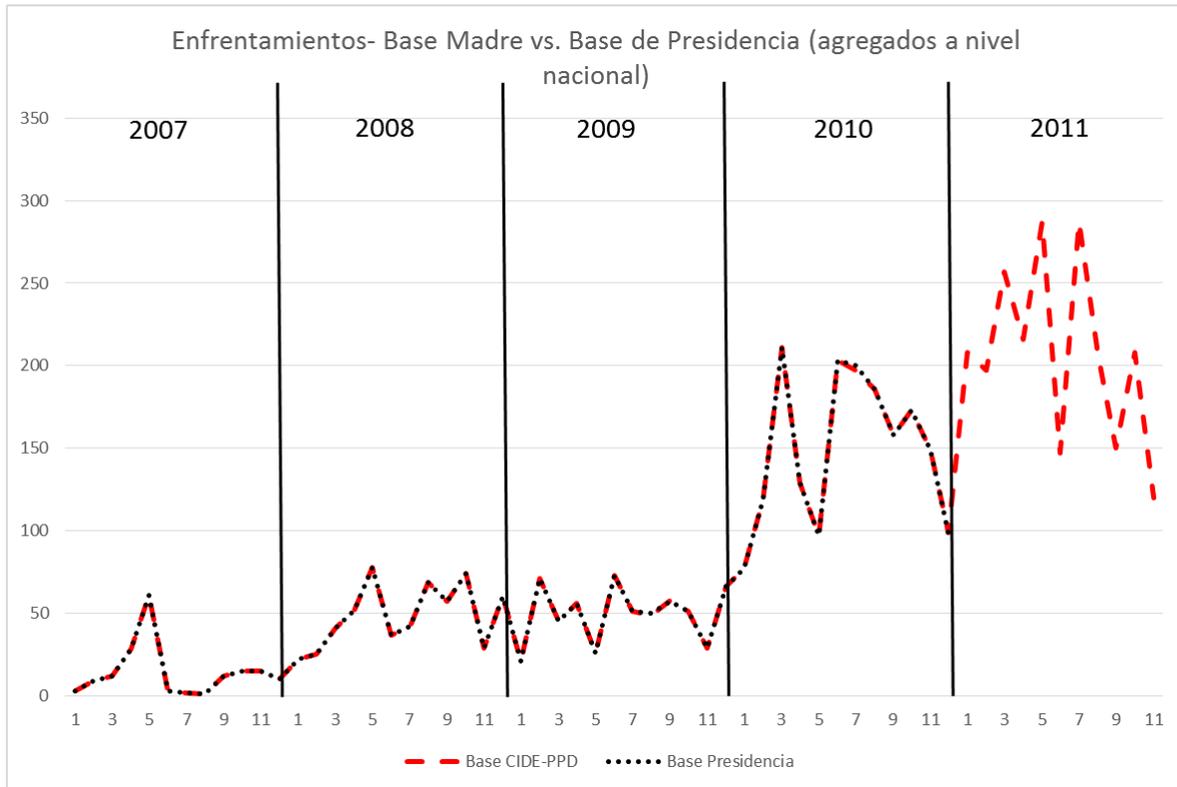
La Base Madre (y, en consecuencia, la Base CIDE-PPD) se compone de tres archivos distintos, a saber: (i) Agresiones, (ii) Enfrentamientos y (iii) Ejecuciones. Los títulos de estas tres bases llamaban la atención, pues coincidían con las categorías utilizadas por la “Base de datos de fallecimientos ocurridos por presunta rivalidad delincuencia” (en adelante, Base Presidencia), publicada en el portal de Presidencia (www.presidencia.gob.mx/voceria_seguridad/) en enero de 2011¹. En consecuencia, fue allí donde comenzamos nuestra indagatoria, comparando la información publicada por Presidencia con la de la Base Madre. La Base Presidencia difiere de la Base Madre en la medida en que en la primera se registran volúmenes agregados de fallecimientos, mientras que la segunda contiene un registro individual para cada evento. Sin embargo, al agregar los homicidios registrados en los eventos de la Base Madre fue posible comparar los volúmenes totales de homicidios en un registro temporal. Al hacer esta comparación, encontramos que —al menos a nivel nacional— el número de homicidios, mes a mes, coincide casi exactamente con los que reporta la Base Presidencia. Las gráficas 1, 2 y 3 comparan cada una de las categorías — **Agresiones, Enfrentamientos y Ejecuciones**— de ambas bases.

Gráfica 1: Comparación del número de homicidios por mes en la categoría de **Agresiones** incluidos en la Base Madre y en la Base Presidencia

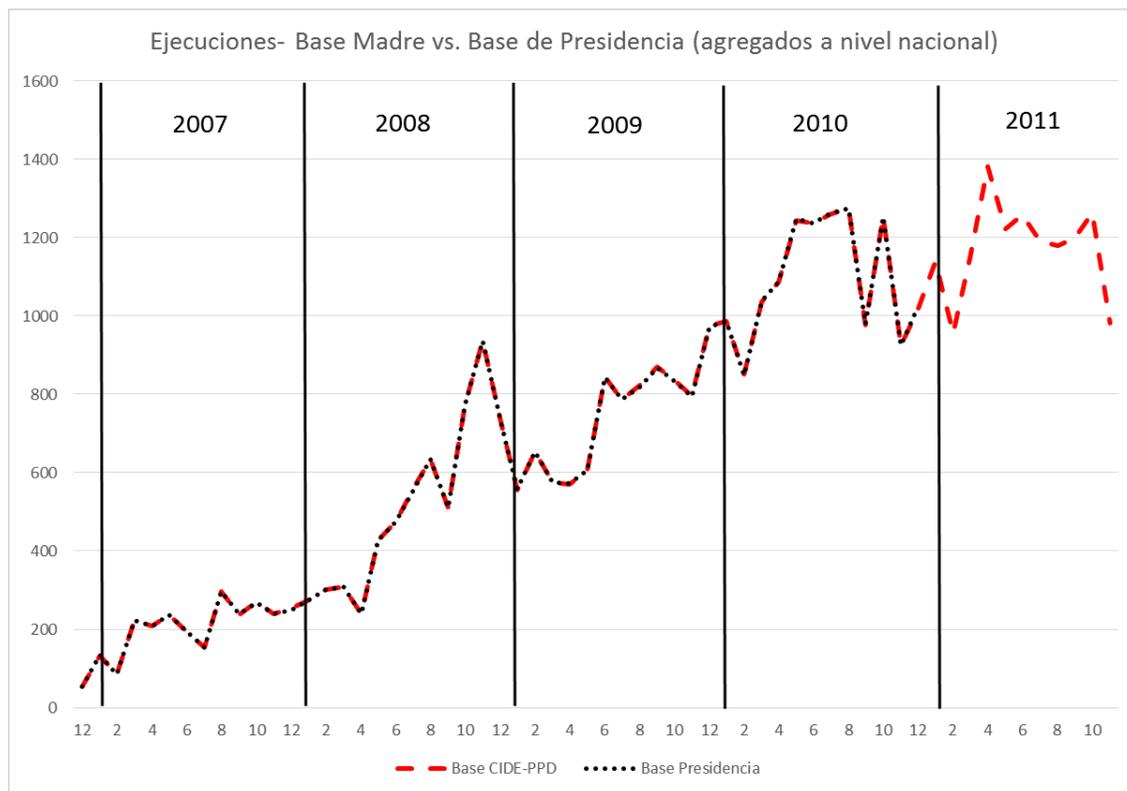


¹ La administración del Presidente Enrique Peña Nieto eliminó el archivo correspondiente del portal, por lo que ya no es consultable en línea. Véase nota 3.

Gráfica 2: Comparación del número de homicidios por mes en la categoría de **Enfrentamientos** incluidos en la Base Madre y en la Base Presidencia



Gráfica 3: Comparación del número de homicidios por mes en la categoría de **Ejecuciones** incluidos en la Base Madre y en la Base Presidencia



Como se puede apreciar, la principal diferencia entre ambas bases consiste en que la Base Presidencia contiene registros hasta diciembre de 2010, mientras que la Base Madre cubre hasta noviembre del siguiente año. El Anexo 1 muestra las diferencias mes con mes para cada una de las tres categorías. En el caso de las **Agresiones** sólo existe una discrepancia entre las dos bases, correspondiente a junio de 2010, mes en el cual la Base Presidencia reporta 34 fallecimientos, mientras que la Base Madre reporta 31. En el caso de los **Enfrentamientos** se encontraron cuatro discrepancias: dos en agosto y septiembre de 2009, y otras dos en julio y diciembre de 2010. En todas ellas la discrepancia es de menos de cuatro homicidios y el volumen de fallecimientos es considerablemente mayor. En el Cuadro 1 se describen estas cuatro discrepancias.

Cuadro 1: Discrepancias en el número de fallecimientos registrados como **Enfrentamientos** en ambas bases

	Base Presidencia	Base CIDE-PPD	Diferencia
Agosto 2009	50	49	1
Septiembre 2009	57	58	1
Julio 2010	200	197	3
Diciembre 2010	98	99	1

En el caso de **Ejecuciones**, la más voluminosa de las categorías, las discrepancias empiezan en 2007, pero todas ellas, salvo dos, son menores a cuatro homicidios². En agosto de 2009 la diferencia es de cuatro homicidios y en diciembre de 2010 es de ocho. El Anexo I precisa estas variaciones mes por mes, pero vale la pena señalar que el número total de discrepancias registradas hasta diciembre 2010 (periodo en el que las dos bases cuentan con información) es de 56. Sin embargo, al hacer la suma de los totales de diciembre de 2006 a diciembre de 2010, el agregado es exactamente igual para las dos bases: 30,812 homicidios.

Aun cuando la semejanza entre ambas bases nos llevó a pensar con un alto grado de certeza que la Base Presidencia podría haber sido extraída de la Base Madre, pues las dos reportaban prácticamente el mismo número de fallecimientos, esta coincidencia no bastaba para considerar válidos los registros de la Base Madre. Esta última contenía más variables y más información que la Base Presidencia. Por ello realizamos un ejercicio sistemático e independiente de validación de los eventos de la Base Madre que describiremos más adelante, aunque podemos anticipar que dicho ejercicio resultó satisfactorio, por lo que consideramos que efectivamente la información de la Base Madre proviene de registros administrativos de las autoridades. Habiendo validado independientemente la Base Madre y encontrado tanta proximidad entre lo reportado por ésta y por la Base Presidencia concluimos que, efectivamente, la Base Madre es la fuente original de la cual se extrajo la información que se hizo pública en la Base Presidencia. Por ello, utilizamos la Guía metodológica publicada con la Base Presidencia³ como un referente para entender cómo fue construida la Base Madre y cuáles son sus limitaciones y su utilidad.

Este documento, y la Base CIDE-PPD que acompaña, es una contribución relevante para entender la violencia relacionada con el crimen organizado en el contexto de la “guerra contra las drogas” lanzada durante el sexenio de Felipe Calderón. Es la primera vez que la información aquí contenida es analizada y publicada. Los datos que ofrecemos permitirán a los investigadores responder preguntas referentes al papel que ha desempeñado en este periodo la coacción estatal en el incremento de la violencia; la explosión de violencia resultante de los conflictos dentro de y entre organizaciones de la delincuencia organizada; la manera en que la violencia funge como un medio de comunicación entre los grupos delincuenciales, con las fuerzas del Estado y con el público en general; y las prácticas de las fuerzas estatales al enfrentarse o detener a presuntos delincuentes; entre otras cuestiones. En síntesis, con la Base CIDE-PPD que hacemos pública, tanto la academia como las personas interesadas en esta temática tendrán un insumo de enorme importancia para plantear nuevas preguntas o verificar respuestas ya formuladas acerca de la violencia generada en torno a la guerra contra las drogas en México durante el periodo 2006-2011.

² Usamos el término “homicidios” al referirnos a las **Ejecuciones** y no “fallecimientos”, pues —a diferencia de los eventos registrados como **Enfrentamientos** y **Agresiones**, en los que muchas de las muertes se presentan como resultado del uso de la fuerza pública en legítima defensa por parte de las autoridades—, en **Ejecuciones** no parece haber duda de que se trata de instancias en que se cometió el delito de homicidio.

³ Tanto la Base Presidencia como su Guía Metodológica fueron retiradas del portal de la Presidencia al inicio del sexenio de Enrique Peña Nieto. Ambas se encuentran en el archivo de los autores.

Este Cuaderno se organiza de la siguiente manera: La sección II sintetiza el contexto en el que se generó la Base Presidencia (asumimos que, en consecuencia, también la Base Madre) y se hizo parcialmente pública, así como las razones que en su momento se ofrecieron para suspender su publicación y finalmente, ya en la administración de Enrique Peña Nieto, para retirarla del portal de internet de la Presidencia de la República. La sección III describe las tres categorías que conforman la Base Madre según lo indica la Guía metodológica de la Base Presidencia. La sección IV describe los procesos de validación de la Base Madre. La sección V describe en detalle la Base CIDE-PPD y su proceso de codificación. Adicionalmente se presenta un análisis temporal y espacial de las principales variables incluidas en cada categoría. La sección VI señala y explora las posibles limitaciones de la Base CIDE-PPD, específicamente aquellas derivadas de la Base Madre, así como la información faltante dentro de los registros que establece. La sección VII concluye ofreciendo algunas preguntas de investigación que podrían responderse mediante el uso de la Base CIDE-PPD y subraya la importancia de hacer pública esta información. Finalmente, el Anexo 1 presenta las discrepancias entre la Base Madre y la Base CIDE-PPD, mientras que el Anexo 2 reproduce textualmente la Guía metodológica de la Base Presidencia (en adelante, Guía Metodológica o Anexo 2).

II. Antecedentes de la Base Presidencia (y, por inferencia, de la Base Madre)

En 2009 el Centro de Investigación y Seguridad Nacional (CISEN), en colaboración con la Procuraduría General de la República (PGR), convocó a un grupo de trabajo a fin de recabar y clasificar la información existente para crear una base de datos de “fallecimientos ocurridos por presunta rivalidad delincuencia” (Guía metodológica). En enero de 2011, la información generada por ese grupo fue publicada por la Oficina de la Presidencia en respuesta al compromiso adoptado por el gobierno de publicar la información relacionada con su estrategia de seguridad y la “lucha contra el crimen organizado”, en el contexto de los Diálogos por la Seguridad convocados por el Presidente Calderón al inicio de su administración (Presidencia de la República, 2011).

Desde la publicación de la Base Presidencia en enero de 2011, la PGR fue criticada porque del total de homicidios relacionados por el crimen organizado que se reportaba en esa Base, sólo el 4.44 por ciento (1,778) tenía abierta una averiguación previa. Según el Código Federal de Procedimientos Penales, todo fallecimiento susceptible de clasificarse como homicidio debe resultar en la apertura de una averiguación previa, ya sea por parte del ministerio público federal o alguno de los estatales. Al no haber averiguación previa de la enorme mayoría de los fallecimientos reportados, la pregunta obligada, en consecuencia, era ¿con base en qué investigación reportaba Presidencia que los fallecimientos eran homicidios y, más aún, específicamente relacionados con la delincuencia organizada? Tras diversas solicitudes de información al respecto, el gobierno retiró temporalmente la base de datos del portal de internet de Presidencia hasta no realizar una revalidación de la información. Por su parte, la

PGR argumentó que correspondía a los ministerios públicos locales investigar el grueso de los homicidios y que ellos deberían validar la información (Proceso, 2012).

En enero de 2012, la PGR publicó una base de datos sobre homicidios relacionados con actividad criminal que cubría hasta el 20 de septiembre de 2011 y que usaba categorías distintas a la Base Presidencia (PGR, 2012). En mayo de 2012, en el contexto de la 27ª Sesión Plenaria de la Conferencia Nacional de Procuradores de Justicia, los procuradores locales consideraron que esta nueva base no era confiable y “decidieron reconstruirla con la misma metodología para tener certeza de las cifras con datos verificados por cada dependencia” (Vanguardia, 2012). A la fecha no se ha publicado una base de datos verificada por las procuradurías locales.

III. Análisis de la Guía metodológica de la Base Presidencia

Esta sección sintetiza lo que dice la Guía metodológica de la Base Presidencia. Según este documento “con finalidad de dar seguimiento al fenómeno de la violencia ocurrida en el contexto de la rivalidad entre las organizaciones criminales, principalmente dedicadas al tráfico de drogas, se instruyó a las dependencias del gobierno federal con competencia en la materia, [a] desarrollar un registro sistematizado y confiable de los homicidios generados en este contexto.” El grupo se compuso por representantes de las Secretarías de Defensa Nacional, Marina, Seguridad Pública Federal y Gobernación, a través del CISEN, y de la Procuraduría General de la República (PGR).

Para esos efectos, el grupo estableció la categoría de “fallecimientos ocurridos por presunta rivalidad delincencial”, sin pretender que tuviera sustento legal sino con propósitos meramente analíticos. Este universo se conformaría de tres bases de datos: una en la que se registrarían las “ejecuciones”, otra en la que se inscribirían los “enfrentamientos” (contra la autoridad y entre delincuentes) y una tercera que habría de asentar las “agresiones” hacia la autoridad. El documento precisaba que los registros debían corresponder a investigaciones en curso, por lo que eventos que inicialmente podrían parecer vinculados a rivalidad criminal, con el paso del tiempo podrían ser retirados si así lo exigía el resultado de las investigaciones.⁴ Esto implicaría que la base de datos estaba pensada para ser dinámica a lo largo del tiempo, y que la información contenida en ella en un momento determinado no fuese definitiva, pues existía siempre la posibilidad de que una investigación posterior obligara a incluir o retirar registros⁵.

Con estas limitaciones, el documento precisa las categorías utilizadas. A continuación reproducimos la información que el documento ofrece al respecto:

⁴ La Guía metodológica no especificaba qué tipo de investigaciones, pero que en todo caso, las investigaciones penales formales (esto es, las averiguaciones previas) fueron solo un mínimo de los casos registrados, tal como se comentó anteriormente en esta sección.

⁵ Esto puede explicar, por ejemplo, las discrepancias mensuales entre la Base Presidencia y la Base Madre, descritas en la sección anterior, que sin embargo arrojan resultados agregados idénticos.

Ejecuciones

Bajo el rubro **Ejecuciones** se incluye todo homicidio doloso “cuya víctima y/o victimario es presumiblemente miembro de algún grupo criminal. No es resultado de un enfrentamiento ni de una agresión, tal como se define en el presente documento. Tampoco presupone la participación de autoridad alguna”. Además, precisa dos características que pueden tener estos eventos: (i) que presenten “extrema violencia”, esto es “que la víctima se localiza decapitada, desmembrada, mutilada y/o calcinada”; (ii) multiplicidad, esto es “aquellos eventos de muerte violenta que implican más de dos víctimas.” A mayor abundamiento, la Guía metodológica precisa qué características debe reunir una muerte violenta para ser incluida bajo este rubro⁶.

Enfrentamientos

Por **Enfrentamientos**, el documento define los actos violentos perpetrados por presuntos criminales en contra de autoridades, víctimas o eventos que perturban el orden público y que se realizan “mediante el uso de armas de fuego y equipo militar”, eventos en los que las propias fuerzas públicas hacen uso de las armas de fuego (en contra de presuntos delincuentes, se entiende), o choques entre grupos delincuenciales o dentro de éstos. Igual que en el caso de **Ejecuciones**, el documento establece una serie de criterios, no siempre claros y en ocasiones redundantes, para precisar estos supuestos⁷.

Agresiones

Ataques de organizaciones delictivas en contra de instalaciones gubernamentales o bien contra funcionarios públicos encargados de la seguridad, sin que la autoridad “tenga posibilidades de responder de forma armada”. Estos eventos se refieren a acciones “dirigidas a personas u objetivos específicos”. Nuevamente, el documento precisa supuestos poco claros.

Es importante destacar que las definiciones y complejas baterías de “características” que establece el documento parecen más criterios que orientan la selección de casos —o bien que racionalizan una selección ya dada— y difícilmente pueden fungir como parámetros discretos que permitan clasificar un evento determinado inequívocamente en uno u otro rubro (o excluirlo). Como ya se comentó, el Anexo 2 reproduce textualmente la Guía metodológica y cada lector podrá formar su opinión al respecto. Por nuestra parte, basta con aclarar que creemos que los criterios son útiles solo si se tiene mucha más información o conocimiento de los eventos, información que no está explicitada en los criterios. En cualquier caso, consideramos importante tener estos criterios presentes pues informan o ilustran las razones —buenas o malas— que tuvieron las autoridades que seleccionaron los eventos y los incluyeron en la Base Madre. Es importante que se critique la metodología con la que fue construida la Base si se pretende revisarla o mejorarla. Pero para quienes, sin ser responsables de la recolección de datos, queremos usar la Base para entender lo ocurrido en México durante ese periodo, los eventos resultan interesantes con independencia de los criterios seguidos para incluir y clasificar los eventos, por el hecho mismo de haber sido seleccionados por las autoridades para conformar una base que les permitiera registrar los sucesos violetos de la

⁶ Ver Anexo 2.

⁷ Ver Anexo 2.

“guerra contra las drogas” del sexenio de Felipe Calderón. Como mínimo, la Base Madre refleja el registro puntual de la “guerra” que llevaba a cabo el gobierno federal.

La Guía metodológica estuvo disponible al público durante un periodo determinado y la base a la que acompañaba ha sido utilizada en distintas investigaciones ya publicadas y consideradas aportaciones de relevancia al estudio de la violencia que México ha vivido a partir del lanzamiento de la “guerra contra las drogas”. Sin embargo, aunque esas investigaciones retomaban los rubros de las categorías referidas, solían describirlas en formas distintas y divergentes de las explicitadas en la Guía metodológica. Por ejemplo, Ríos (2013), en uno de los textos más citados para explicar la reciente violencia en México, define confrontaciones, ejecuciones y agresiones de forma similar a la establecida en la Guía metodológica, pero obviando muchas de las contradicciones, indeterminaciones y redundancias de ésta. No obstante, utiliza enfrentamientos y agresiones como *proxies* de competencia entre grupos delincuenciales y aplicación de la ley, respectivamente, cuando los criterios establecidos en la Guía metodológica no permiten una equiparación así de sencilla. Por ejemplo, los **Enfrentamientos**, según la Guía, pueden suscitarse entre delincuentes o entre delincuentes y autoridades, por lo que no todos deben considerarse eventos que reflejan competencia entre grupos delincuenciales.

Calderón *et al.* (2015), por su parte, dan definiciones más sencillas de las tres categorías: las **Ejecuciones** se definen como homicidios resultantes de rivalidad entre grupos delincuenciales, las **Agresiones** como ataques a autoridades, y los **Enfrentamientos** como muertes suscitadas en el contexto de una acción de las fuerzas de seguridad (sin hacer distinción entre los dos tipos de enfrentamientos que se incluyen en la Base). A su vez, aunque utiliza dos y no tres categorías, Dell (2014) reporta usar datos Consejo Nacional de Seguridad que parece corresponden a las **Ejecuciones** de la Base Presidencia, por una parte; y el agregado de **Ejecuciones** y **Agresiones** por la otra.⁸

Los ejemplos aquí citados se ofrecen, no como críticas a los trabajos, sino como evidencia de las limitantes de la Base Presidencia para poder hacer un análisis de la violencia en México durante ese periodo. Al no tener más información de los registros excepto el número de homicidios, es difícil especular sobre las dinámicas de la violencia, tales como el papel que juegan las fuerzas públicas, los motivos de los homicidios registrados, o bien el perfil de las víctimas (delincuentes, autoridades o terceros). La Base CIDE-PPD ofrece mucha más información de cada registro, cuyo análisis y cruce quizá permita sofisticar los trabajos ya realizados y ofrecer mayor profundidad en la comprensión de los sucesos que documenta.

⁸ Son varios los indicios de que Dell usa la Base Presidencia o bien una versión de la Base Madre. Para empezar, cuando uno analiza los eventos contenidos en **Agresiones** y **Ejecuciones** parece que la distinción terminó por ser arbitraria y que podrían haberse agregado estas dos bases, como parece hacerlo ella. Además, la última fecha de sus datos —octubre de 2011— es próxima a la fecha en que concluye nuestra versión de la Base Madre. Finalmente, la fuente confidencial que dice usar la refiere a un comité con representantes de todos los ministerios que participan en el Consejo Nacional de Seguridad Pública; la periodicidad de las reuniones de ese comité y la posibilidad de reclasificar eventos coincide con la información con la que contamos acerca de la elaboración de la Base Presidencia.

IV. Validación de la Base Madre

El Programa de Política de Drogas del CIDE, en colaboración con CentroGeo, sometió la Base Madre a dos procesos distintos. El primero fue la validación de sus registros. El segundo fue la codificación de algunos de los campos contenidos en dicha Base Madre, a fin de garantizar la seguridad de las personas involucradas en los eventos registrados. La validación se realizó contrastando la información contenida en los registros individuales con información pública accesible en internet, principalmente noticias y otras fuentes abiertas. A continuación se describe el primero de estos procesos. El segundo se describirá en la sección V.

De los 36,066 registros contenidos en las tres bases de datos que conforman la Base Madre, seleccionamos las variables más consistentemente reportadas: nombre de la víctima —cuando lo había—, ubicación, fecha y número de fallecidos, lesionados y detenidos registrados para el evento. Esta información fue utilizada para realizar búsquedas automáticas de información pública en internet, a través del Autómata Geointeligente en Internet (AGEI) que CentroGeo desarrolla como parte de la construcción del Laboratorio Nacional de GeoInteligencia Territorial. En promedio, se encontraron veinte enlaces relacionados con los términos de búsqueda de cada evento considerado. Estos enlaces fueron utilizados para generar una base de datos de sitios web públicos, que fueron cotejados por un equipo de validadores por medio de una herramienta diseñada para ese fin, llamada SAVE (Sistema de Análisis y Validación de Eventos). De una forma sencilla e intuitiva, el SAVE permitía a los validadores cotejar información de la Base Madre con información encontrada en los sitios web públicos, resaltando automáticamente los términos coincidentes de mayor relevancia. En todos los casos, si todas las variables consideradas podían corroborarse en alguno de tales sitios, sin encontrar en ellos información contradictoria, el evento se calificaba como validado.

Dentro del proceso de validación, cada evento podía ser calificado de tres formas distintas:

- (i) Se validaba completamente, si todas las variables de la Base coincidían con la información de los sitios web;
- (ii) Si algunas, pero no todas, las variables coincidían, se evaluaba la discrepancia conforme a parámetros preestablecidos y podía entonces o bien validarse el evento, o bien introducir una nota al evento sin validarlo. Por ejemplo, si la fecha discrepaba por uno o dos días, el evento se validaba; si la discrepancia recaía sobre cualquier otra variable, como el número de fallecimientos, se consideraba parcialmente validado y se hacía una anotación para análisis posterior;
- (iii) Si no se encontraba el evento en las fuentes abiertas de internet, no se validaba.

La tercera posibilidad, cuando el evento no encuentra referente alguno en internet, nos obliga a reflexionar sobre posibles causas. No todos los eventos relacionados con la delincuencia organizada son reportados por alguna fuente abierta. Existen varios factores conocidos que afectan la libertad de prensa para reportar un evento. Por ejemplo, la delincuencia organizada puede presionar a los medios de comunicación para que no reporten eventos violentos en

lugares específicos. O un evento puede no ser interesante para la prensa por no ser “suficientemente violento”, es decir: en la medida en que la guerra contra el crimen organizado avanza, y los modos de muerte usados por los grupos criminales escalan, los eventos se vuelven menos interesantes en la medida en que son menos violentos. O quizá el mismo día suceden eventos más significativos, que hacen que el evento en cuestión no sea noticia. Cualquiera que sea la razón, existen diferencias temporales y espaciales en el proceso de validación que serán analizadas en esta sección y que deberían considerarse para realizar análisis sobre el comportamiento del crimen organizado.

De acuerdo con un reporte presentado en 2010 por ARTICLE 19 y el Centro Nacional de Comunicación Social (Cencos), solo durante 2009 se registraron 244 agresiones a la libertad de expresión en México, las cuales incluyeron asesinatos (4.51 por ciento), intimidación (19.26 por ciento), agresión física material (44.67 por ciento), detenciones (9.43 por ciento), amenazas (19.26 por ciento) y desapariciones (0.82 por ciento). De este total, se presume que en 65.57 por ciento de los casos los funcionarios públicos fueron los agresores, y en 6.15 por ciento el crimen organizado⁹. Por otro lado, puede que un evento no haya aparecido en prensa porque sucedió en sitios aislados o porque el mismo día existieron otros eventos que fueron considerados más importantes para reportar. Independiente de cuál sea el motivo por el cual no se reportó el evento, el hecho de que la Base Madre contenga más eventos de los registrados en la prensa indica que su cobertura es más extensa que las de otras bases de datos construidas solamente con fuentes abiertas.

El Cuadro 2 muestra el número de eventos validados, al igual que el número total de sitios web revisados en el proceso de validación. En este cuadro se consideran tanto los eventos validados directamente, como aquellos que fueron encontrados con discrepancias menores de acuerdo con los parámetros preestablecidos:

Cuadro 2: Total de eventos, número de eventos encontrados en internet, número de eventos validados y porcentaje de validación para cada una de las categorías de la Base Madre

Categoría	Total de eventos	Total de eventos encontrados en Google	Total de eventos validados	Porcentaje de validación
Agresiones	1,561	1,561	457	29%
Enfrentamientos	3,835	3,827	895	23%
Ejecuciones	30,982	30,652	15,012	48%

En el proceso se descargaron 720,297 sitios web (20 por evento en promedio) y se revisaron 469,820 sitios. Con esta revisión se pudieron validar 16,364 eventos; se encontraron 618 eventos en internet, pero con información incompleta (del nombre de la víctima, la fecha o el lugar de los hechos, etc.), y 452,638 sitios no estaban relacionados en lo absoluto con el evento en cuestión. Estas cifras se resumen en el Cuadro 3 a continuación:

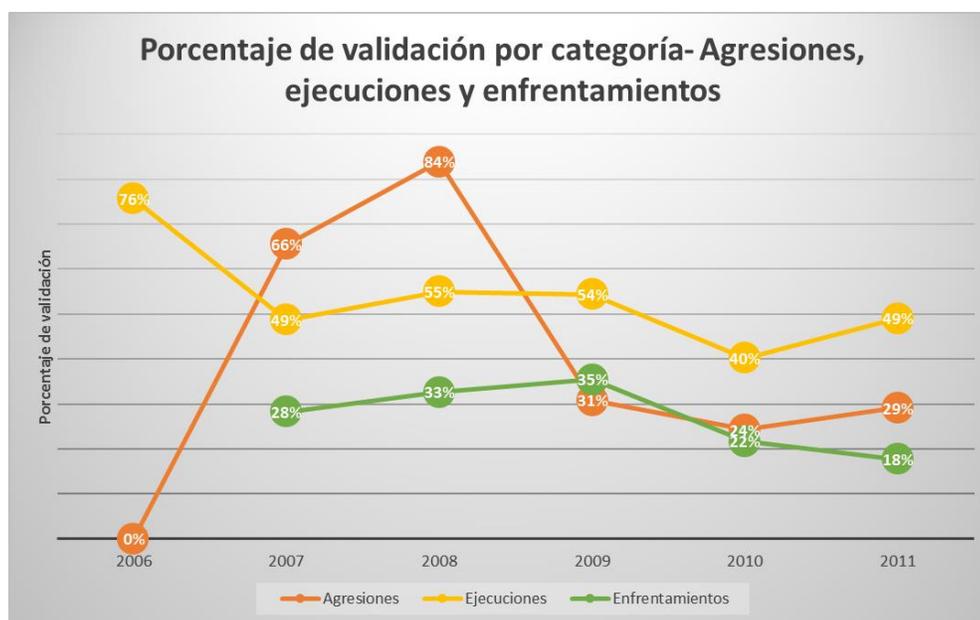
⁹ Según el mismo reporte, de 160 agresiones a la prensa en las que se presume algún funcionario público como responsable, 45 corresponden a alguna fuerza pública federal, 58 estatal y 57 municipal.

Cuadro 3: Categorización de los *links* usados en el proceso de validación

Categorización de los <i>links</i>	Totales
Número de <i>links</i> descargados	720,297
Número de <i>links</i> revisados	469,820
Número de <i>links</i> que sirvieron para validar eventos	16,364
Número de <i>links</i> parcialmente validados (con anotaciones acerca de las discrepancias)	618
Número de <i>links</i> no relacionados con los eventos	452,638

A nivel nacional, la proporción de eventos validados varía dependiendo del año y de la categoría analizada (Agresiones, Enfrentamientos y Ejecuciones). La Gráfica 4 muestra el comportamiento de la validación tomando en cuenta categoría y tiempos.

Gráfica 4: Porcentaje de validación por categoría: Agresiones, Ejecuciones y Enfrentamientos

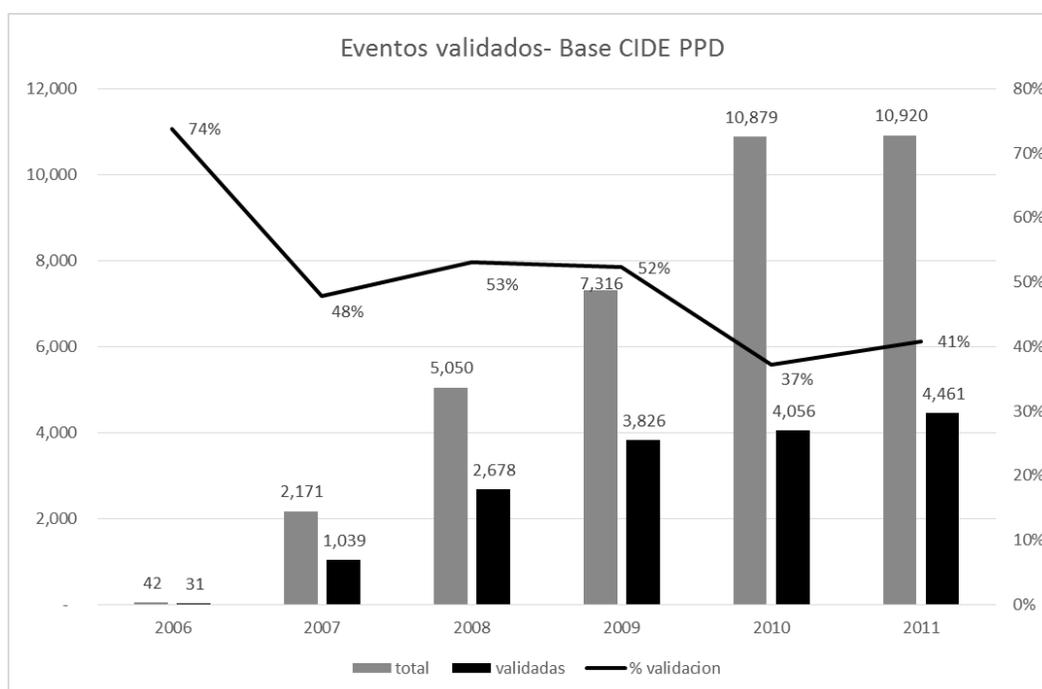


En promedio, la categoría más validada fue la de **Ejecuciones** (54 por ciento), seguida de **Agresiones** (39 por ciento) y **Enfrentamientos** (27 por ciento). La mayor proporción se registra en **Agresiones** del año 2008, con un porcentaje de validación de 84 por ciento: de un total de 31 eventos, se validaron 26. En promedio, la categoría que mejor se validó en todos los años fue la de **Ejecuciones** (con excepción de 2008, año en el que se validaron más **Agresiones** que **Ejecuciones**). Sin embargo, se puede ver que el porcentaje de validación va cayendo a medida que va pasando el tiempo. En el caso de las **Ejecuciones**, el porcentaje de validación fue de 76 por ciento en diciembre de 2006, y para 2011 esta cifra fue de 49 por ciento. A pesar de que no existe un factor único para explicar este comportamiento, varios sucesos pudieron haber afectado el proceso de validación, entre ellos: las crecientes agresiones a la prensa en las diferentes entidades federativas; el crecimiento acelerado de la violencia (lo

que hizo que las muertes *menos* violentas dejaran de ser noticia), o los pactos firmados por la prensa para no reportar todas las acciones relacionadas con el crimen organizado.¹⁰

La Gráfica 5 muestra el total de eventos en cada uno de los años, el total de eventos validados y su porcentaje de validación (en términos agregados para las tres categorías). Como se puede observar, en la medida que evolucionaba la guerra contra las drogas, el número de eventos se incrementaba, al igual que el número de eventos validados. Sin embargo, la tasa de crecimiento de los eventos totales es mayor a la de eventos validados, razón por la cual el porcentaje de validación es cada vez menor. Por ejemplo, de 2,171 eventos observados en 2007, 1,039 fueron validados (correspondientes a un porcentaje de validación de 48 por ciento). En 2011, el número de eventos totales se incrementó a 10,920; y de estos, 4,461 fueron validados, lo que corresponde a un porcentaje de validación de 41 por ciento.

Gráfica 5: Eventos validados y porcentaje de validación por año



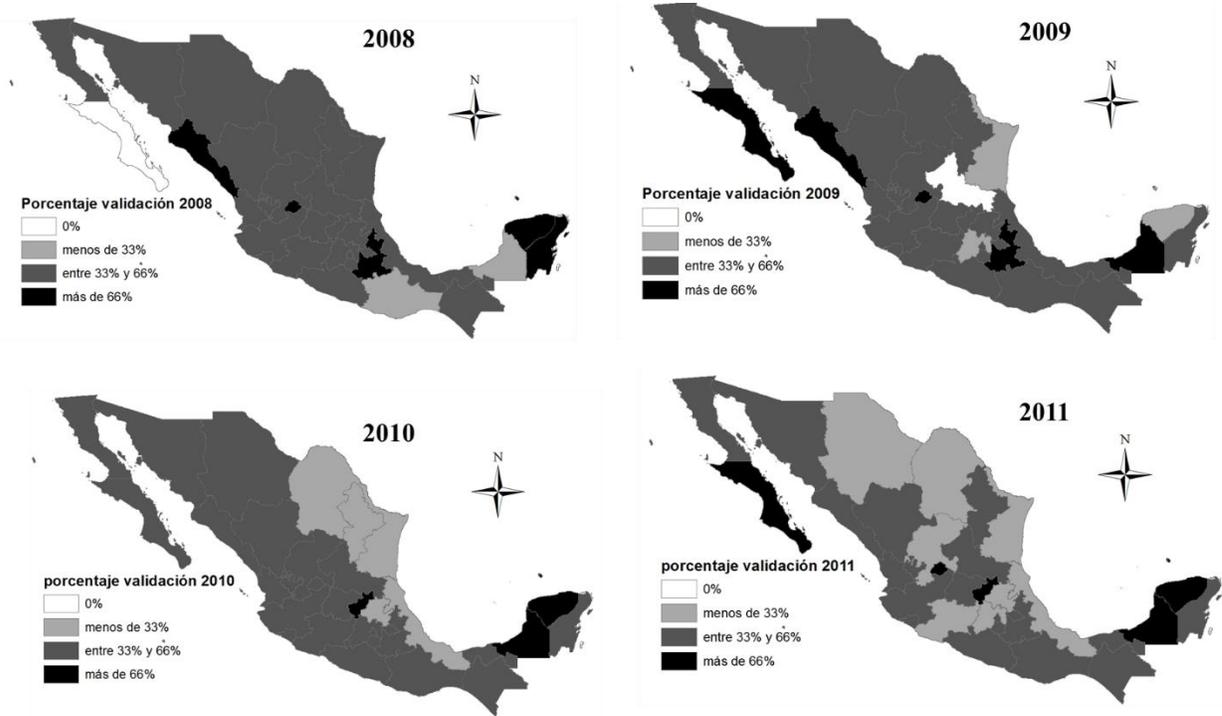
Las diferencias espaciales se pueden observar en los mapas de la Gráfica 6. Se presenta un mapa para cada año con los porcentajes totales de validación por entidad federativa. Las diferencias son bastantes significativas, tanto temporal como espacialmente. En 2008 casi todos los estados tuvieron un porcentaje de validación entre 33 y 66 por ciento. Mayores porcentajes se observaron en Aguascalientes (90 por ciento), Puebla (76 por ciento), Quintana Roo (67 por ciento), Sinaloa (72 por ciento) y Yucatán (78 por ciento). En 2010, ya todos los estados en la cuenca del Golfo de México presentaron porcentajes de validación menores al 33 por ciento, y en 2011 esta tendencia se extendió también a los estados localizados en el norte

¹⁰ El pacto más importante, firmado en marzo de 2011, fue el llamado “Acuerdo para la Cobertura Informativa de la Violencia”, en el que los principales medios de comunicación se comprometían a “no interferir en el combate a la delincuencia, dimensionar adecuadamente la información y no convertirse en voceros involuntarios” de los criminales (*La Jornada*, 2011).

del país, incluidos Coahuila y Chihuahua, y a otros estados que sufrieron elevados incrementos de violencia como Michoacán, el estado de México y Zacatecas. A pesar de que la base de datos no ofrece evidencia específica sobre la razón de este comportamiento, al observar la evolución temporal lo que sí parece suceder es que, a medida que se incrementa la violencia, el porcentaje de eventos que se reportan en los medios es cada vez menor; y este comportamiento es más fuerte en zonas con mayor violencia relacionada con el crimen organizado.

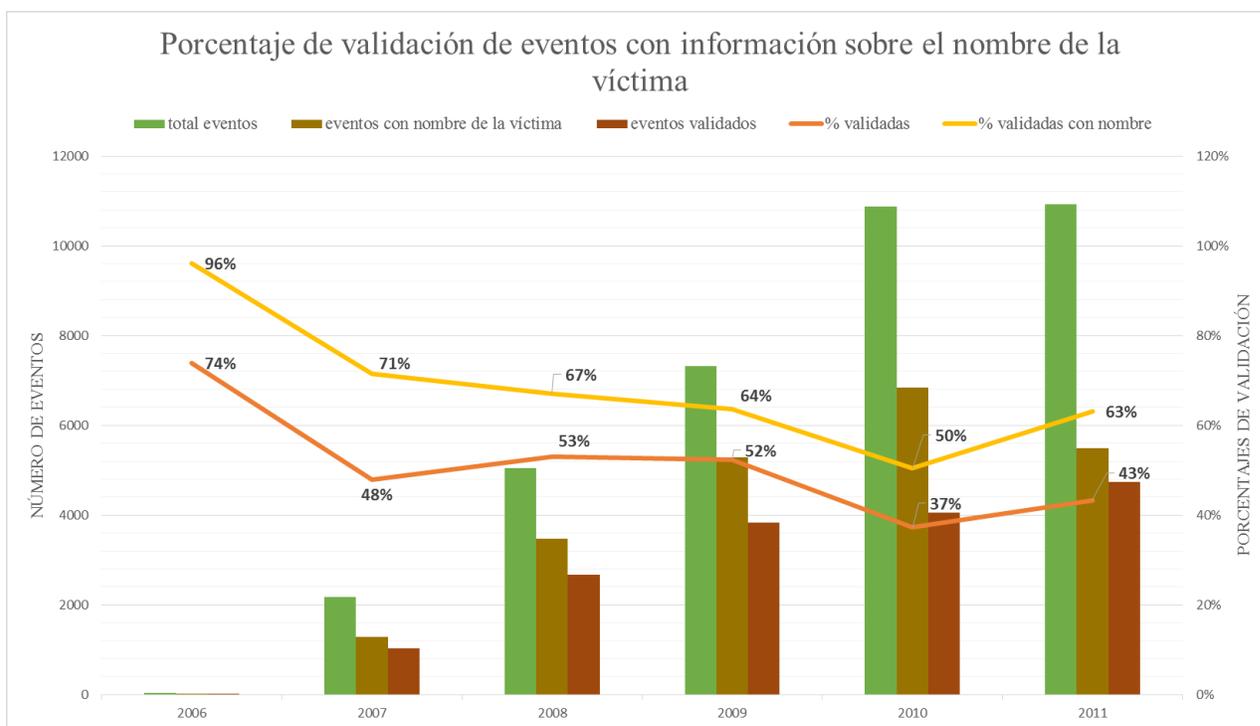
De acuerdo con ARTICLE 19 (2012), las entidades federativas que presentaron un mayor número de periodistas y medios agredidos fueron Veracruz (donde las agresiones se incrementaron en 200 por ciento), Distrito Federal, Chihuahua (el estado con más periodistas asesinados desde 2000 después de Tamaulipas), Coahuila y Oaxaca. Con excepción del Distrito Federal y de Oaxaca, estos estados se incluyeron entre aquellos en los que el porcentaje de validación se encuentra por debajo de 33 por ciento: Veracruz, con un 27 por ciento de validación, Chihuahua con 32 por ciento, y Coahuila con 24 por ciento. De acuerdo con la misma fuente, el narco también ha utilizado a los medios locales para difundir sus intereses, sesgando la información encontrada en fuentes abiertas y, consecuentemente, la validación aquí descrita. Por ejemplo, en Tamaulipas los medios empezaron a recibir despachos informativos y redacciones por parte del crimen organizado que debían ser publicados. En Michoacán, un integrante de La Familia Michoacana informó por medios escritos y visibles sus preferencias políticas en contra del partido político en el poder.

Gráfica 6: Porcentajes de validación por entidad federativa y por año



Finalmente, incluimos un análisis del porcentaje de validación por variable utilizada. En las tres categorías, cuando se incluía el nombre de la(s) víctima(s), las probabilidades de validar el evento eran mucho más altas. La Gráfica 7 muestra la relación entre eventos en los que se incluía el nombre de la víctima y el porcentaje de validación de 2006 a 2011. En todos los años, el porcentaje de validación fue más alto para las observaciones en las que se incluía información sobre el nombre de la víctima que para el total de eventos. Considerando solo los eventos en los cuales se incluía el nombre de la víctima, los porcentajes de validación mejoran entre 11 y 24 puntos porcentuales dependiendo del año analizado. Por ejemplo, en 2007, el porcentaje de validación de todos los eventos fue de 48 por ciento; este porcentaje se incrementó a 71 por ciento cuando se consideraron solo los eventos en los que se incluía el nombre de la víctima. En 2011 estas cifras fueron de 43 y 63 por ciento, respectivamente, y se observó una diferencia de 20 puntos porcentuales. En promedio para los cinco años, en 63 por ciento de los eventos se incluyó el nombre de la víctima (2009 fue el año con mayor información; el nombre de la víctima se obtuvo en el 72 por ciento del total de eventos registrados en dicho periodo).

Gráfica 7: Relación entre los eventos en los que se incluía el nombre de la víctima y el porcentaje de validación de 2006 a 2011



V. Codificación de la Base Madre y descripción de la resultante Base CIDE-PPD

Al revisar la información recibida, la principal preocupación del Programa de Política de Drogas fue la gran cantidad de datos sensibles que contenía la Base Madre. Ésta incluía no solo nombres de víctimas, sino también información sensible respecto a la ubicación de los

eventos, los funcionarios involucrados, y amenazas recibidas. Aun retirando los principales campos de la base —descripción de eventos, observaciones de campo, nombres de víctimas, detenidos y heridos— había información en otros campos que podía llegar a ser sensible para los individuos involucrados o sus familiares.

En las discusiones internas sobre qué hacer con la información, se llegó a la conclusión de que habría que generar una base distinta, que contuviera solo los campos relevantes para el análisis de políticas públicas y ciencias sociales, pero que omitiera los datos que pudieran ser personalmente sensibles. Esto implicaba codificar la información contenida en ciertos campos a fin de rescatar parte de ésta en forma de variables y excluir el resto. El equipo del PPD se dio entonces a la tarea de definir las variables que quedarían incluidas en la nueva base (la que aquí denominamos Base CIDE-PPD) para su publicación, y que se generarían a partir de los registros sobre “descripciones” y “observaciones” de los distintos eventos. En consecuencia, la Base CIDE-PPD contiene más variables pero menos información que la Base Madre. Estamos conscientes de que no toda la información relevante para el análisis de políticas públicas y ciencias sociales fue rescatada de la Base Madre al realizar la codificación. Desafortunadamente, el número de personas y perfiles que podíamos involucrar en este proceso era limitado. En un futuro, cuando la información contenida en la Base Madre deje de ser sensible, podrán realizarse nuevos esfuerzos de codificación. Pero por ahora consideramos que el ejercicio realizado aporta nueva información valiosa, sin correr el riesgo de afectar a personas en lo individual. A continuación se explican cuáles fueron estas variables y con base en qué presupuestos se construyeron.

a. Nuevas variables para **Enfrentamientos y Agresiones**

El Cuadro 4 contrasta las variables contenidas en la Base Madre con las que quedaron incluidas en la Base CIDE-PPD.

Cuadro 4: Variables contenidas en la Base Madre y en la Base CIDE-PPD

Base Madre	Base CIDE-PPD
Fecha del evento	ID
Estado	Día
Municipio	Mes
Delegación / colonia	Año
Tramo / carretera	Estado
Corporación policiaca involucrada	Municipio
Presuntos delincuentes involucrados	Personas fallecidas
Detenidos	Militares fallecidos
Ocupación (del detenido)	Marinos fallecidos
Nombre (del detenido)	AFI / Policía Federal Ministerial fallecidos
Personas fallecidas	Policías estatales fallecidos
Militares (fallecidos)	Policías ministeriales fallecidos
Marinos (fallecidos)	Policías municipales fallecidos
Policías Federales (fallecidos)	Agentes del Ministerio Público fallecidos
AFI / Policía Federal Ministerial (fallecidos)	Delincuencia organizada fallecidos
Policías Ministeriales (fallecidos)	Civiles y N / E fallecidos
Policías Municipales (fallecidos)	Personas lesionadas

Agentes del Ministerio Público (fallecidos) Delincuencia organizada (fallecidos) Civiles y N / E (fallecidos) Nombre (del / de los fallecido(s)) Personas lesionadas Militares (lesionados) Marineros (lesionados) Policías Federales (lesionados) AFI / Policía Federal Ministerial (lesionados) Policías Ministeriales (lesionados) Policías Municipales (lesionados) Agentes del Ministerio Público (lesionados) Delincuencia organizada (lesionados) Civiles y N / E (lesionados) Nombres (del / de los lesionado(s)) Armas largas (aseguramientos) Armas cortas (aseguramientos) Cargadores (aseguramientos) Cartuchos (aseguramientos) Vehículos (aseguramientos) Otros aseguramientos Hechos Acta circunstanciada Averiguación previa Delito Observaciones	Militares lesionados Marineros lesionados AFI / Policía Federal Ministerial lesionados Policías estatales lesionados Policías ministeriales lesionados Policías municipales lesionados Agentes del Ministerio Público lesionados Delincuencia organizada lesionados Armas largas aseguradas Armas cortas aseguradas Cargadores asegurados Cartuchos asegurados Vehículos asegurados Acta circunstanciada (sí / no) Averiguación previa (sí / no) Delito (sí / no) Días transcurridos desde día 1 (primer Enf. o Agr.) Prensa (sí / no) Fuego cruzado (sí / no) Elecciones (sí / no) Taxi (sí / no) Drogas (sí / no) Vehículos (sí / no) Cuadros externos Otros aseguramientos Vehículos involucrados Drogas involucradas Tipo de presuntos delincuentes Grupo al que pertenecen los presuntos delincuentes Tipo de detenidos Grupo al que pertenecen los presuntos detenidos Delito(s) asociado(s) al evento Entidad policiaca involucrada
Variables adicionales en la base de datos de Agresiones Instalación oficial Agredido Respuesta Granadas Arma de fuego	Variables adicionales en la base de datos de Agresiones Instalación oficial (sí / no) Agredido (especifica quién fue el agredido) Respuesta estatal (sí / no) Uso de granadas (sí / no) Uso de armas de fuego (sí / no)

Como se puede apreciar, la mayoría de las variables de la Base Madre se mantuvieron, con excepción de las que contenían datos como nombres, descripciones de eventos y observaciones. En el caso de las variables “acta circunstanciada” y “averiguación previa”, cuando había alguna asociada con el evento (una minoría del total), se incluyeron variables dicotómicas que indicaban que existía esa asociación, pero excluían los números que remitían al registro. Variables como las incautaciones o las agencias gubernamentales involucradas en el evento fueron codificadas y movidas a una tabla externa, dado que se trataba de variables no excluyentes que no podían ir en el cuadro principal; esto es, variables que podían tener más de un valor para un mismo evento.

A partir de la información contenida en los campos correspondientes a las descripciones de los eventos y a las observaciones de campo, extrajimos información y la codificamos como variables nuevas de forma automatizada. Para ello utilizamos palabras clave contenidas en los campos mencionados, como se indica en el Cuadro 5 a continuación:

Cuadro 5: Palabras clave utilizadas para la creación de las nuevas variables en la Base CIDE-PPD a partir de la información incluida en el campo de descripciones

Variable	Palabras clave
Prensa	Periódico, periodista, prensa, corresponsal, reportero, redactor, radiodifusora, televisora y televisión
Fuego cruzado	Repelió, intercambio de disparos, respondieron a la agresión, tiroteo y enfrentamiento
Elecciones	Campaña electoral, elecciones, mitin, competencia electoral, elección de candidatos, candidatos a elecciones y candidatos
Taxi	Se encontró la palabra “taxi” asociada al evento.
Drogas	Se encontró la palabra “droga” asociada al evento.
Vehículo	Se encontró la palabra “vehículo” asociada al evento.

En cuanto a las tablas externas, se crearon diccionarios específicos a fin de poder involucrar todas las categorías posibles de una variable. Los diccionarios de los cuadros externos se describen en el Cuadro 6 a continuación:

Cuadro 6: Tablas externas incluidos en las bases de **Agresiones** y **Enfrentamientos**

Número y nombre de la tabla	Información contenida
Tabla 1. Otros aseguramientos	<p>En la Base Madre, el campo “otros aseguramientos” consistía en descripciones narrativas. Para la codificación identificamos distintas categorías de objetos confiscados por el gobierno. Los valores del Cuadro son variables dicotómicas que indican la presencia o ausencia de un tipo de objeto determinado. No se incluyeron cantidades porque la información no estaba recabada uniformemente y resultaba imposible homogeneizar las medidas registradas. Las categorías establecidas fueron:</p> <ul style="list-style-type: none"> • Equipos de comunicación • Armamentos • Uniformes • Dinero • Vehículos • Marihuana • Cocaína • Heroína • Otras drogas • Otros aseguramientos
Tabla 2. Vehículos involucrados	<p>En algunas descripciones de los eventos se incluía una descripción más o menos detallada de los vehículos involucrados (tales como modelo, año, lugar de origen de las placas, etc.). Creamos un diccionario a partir de la información más común encontrada acerca de los vehículos:</p>

	<ul style="list-style-type: none"> • Automóvil • Camión • Camioneta • Jeep • Motocicleta • Patrulla • Tanque • No especificado
Tabla 3. Drogas involucradas	<p>Además de los distintos tipos de droga registrados en “otros aseguramientos”, identificamos tipos de drogas incluidos en los campos de “descripciones” y “observaciones”. El diccionario resultante es el siguiente:</p> <ul style="list-style-type: none"> • Cocaína • Crack • Éxtasis • Marihuana • Metanfetamina • LSD • Heroína
Tabla 4. Tipo de presuntos delincuentes	<p>La Base Madre incluía un campo de “presuntos delincuentes” en que se contenía información diversa sobre personas específicas. Con base en esta información se crearon dos diccionarios: uno referente al tipo de presunto delincuente y otro referente al grupo (cuando se mencionaba algún grupo específico). El diccionario sobre el “tipo de presunto delincuente” es el siguiente:</p> <ul style="list-style-type: none"> • Civiles • Fuerza pública • Traficantes de personas • Presuntos delincuentes • Bandas del crimen organizado • Secuestradores • Extorsionadores • Sicarios • Reos • Grupo armado • Narcotraficantes • Narcomenudistas • Traficantes de armas • Grupo identificado (se identificó el grupo en la tabla 5) • Sin determinar
Tabla 5. Grupo al que pertenecen los presuntos delincuentes	<p>En este campo se identificaron 24 grupos criminales diferentes.</p>
Tabla 6. Tipo de detenidos	<p>Al igual que en “tipo de presuntos delincuentes”, el campo de detenidos incluía información genérica y/o específica sobre el/los detenido(s). El diccionario de “tipo de detenido” es el siguiente:</p> <ul style="list-style-type: none"> • Presunto delincuente • Traficantes de personas • Sicarios • Fuerza pública • Comando armado • Narcomenudistas • Narcotraficantes • Secuestradores • Civiles

	<ul style="list-style-type: none"> • Extorsionadores • Escolta • Traficantes de armas • Bandas del crimen organizado • Reos • Grupo identificado (se identificó el grupo en la Tabla externa 7) • Sin determinar
Tabla 7. Grupo al que pertenecen los detenidos	Se identificaron 20 grupos criminales diferentes en este campo.
Tabla 8. Delito(s) asociado(s) al evento	Además de incluir una variable dicotómica en la base principal, se incluyó este Cuadro externo que describe el/los delito(s) asociados al evento con base en el Código Penal Mexicano. Se identificaron 25 delitos diferentes con sus subcategorías.
Tabla 9. Entidad policiaca involucrada	Esta variable ya estaba incluida en la Base Madre. 29 diferentes entidades de la fuerza pública (locales y federales) fueron identificadas. En gran parte de los eventos, hubo participación de más de una entidad.

Existe información adicional, actualmente en proceso de codificación, que no estará disponible en esta primera versión de la Base CIDE-PPD. Sin embargo, una vez que se realicen las nuevas codificaciones, se presentará una actualización de la Base CIDE-PPD que incluirá nuevas variables. Un ejemplo de las nuevas codificaciones que se están llevando a cabo es la identificación de los eventos en los que la fuerza pública tomó la iniciativa de empezar el enfrentamiento. Tal como se describirá más adelante, en la mayoría de los casos, los eventos incluidos en las categorías de **Agresiones** y **Enfrentamientos** son iniciados por presuntos delincuentes o grupos armados. Sin embargo, algunos elementos incluidos en las descripciones permiten identificar ciertas actividades de las fuerzas armadas que hacen pensar que fueron acciones gubernamentales las que originaron el enfrentamiento.

b. Nuevas variables para **Ejecuciones**

El Cuadro 7 compara las variables incluidas en la Base Madre con las variables codificadas en la base CIDE-PPD para la categoría de **Ejecuciones**.

Cuadro 5: Variables incluidas en la Base Madre y en la Base CIDE-PPD en la categoría de **Ejecuciones**

Base Madre	Base CIDE-PPD. Ejecuciones
Fecha	ID
Mes	Día
Estado	Mes
Municipio	Año
Delegación	Estado
Colonia	Municipio
Calles	Ejecutados
Sexo	Ejecutados femeninos
Nombre	Ejecutados masculinos
Edad años	Fuerza pública ejecutada
Policías y/o funcionario público	Cargo
Cargo (no disponible todos los años)	Número de policías
Número de policías (no disponible todos los años)	Averiguación previa y/o acta circunstanciada

Perteneciente a la organización delictiva (no disponible todos los años) Organización delictiva que cometió el delito (no disponible todos los años) Características de los hechos Mensajes Ejecutados Arma Calibre Modo (no disponible todos los años) Averiguación previa (no disponible todos los años) Observaciones (no disponible todos los años)	Mensaje Tortura Días transcurridos desde el primer evento del año Prensa Fuego cruzado Elecciones Corrupción Taxi Drogas Vehículos Tablas externas Vehículos involucrados Drogas involucradas Edad del (de los) ejecutado(s) Tipo de arma(s) utilizada(s) Calibre-munición Modo de muerte Clasificación de los mensajes dejados al lado de los cuerpos Grupos involucrados en el evento
--	--

En caso de existir, la información personal sobre las víctimas de las **Ejecuciones**, o sobre el contexto específico en el que fueron hallados mensajes próximos a un cuerpo, fue retirada de la Base Madre y codificada para ser reintroducida en la Base CIDE-PPD. Algunas otras variables con información incompleta fueron también incorporadas como variables binarias o como categorías. Estas modificaciones se describen a continuación:

- **Mujeres / Hombres ejecutados:** cuando hay más de una ejecución por evento, la variable “sexo” en la Base Madre usualmente no especifica el número de ejecutados de cada sexo. En consecuencia, la única información disponible consiste en la introducción de una variable dicotómica en la que se señala si hubo o no víctimas de cada sexo.
- **Fuerza Pública Ejecutada:** cuando un miembro de un cuerpo de la fuerza pública es la víctima de una ejecución, la Base Madre registra su filiación. Tras limpiar esta variable, creamos un diccionario con 29 categorías distintas en las que pueden clasificarse estas víctimas.
- **Cargo:** Con la excepción de los años 2006 y 2007, la Base Madre incluye información referente al cargo de la víctima de **Ejecuciones**, aunque no en todos los casos existe este registro. Tras limpiar la variable, las posibles categorías de empleo fueron codificadas como sigue: privado-civil, militar, expolicía, exfuncionario, policía, político, funcionario público, custodio, delincuente o extranjero.
- **Averiguación previa (AP) y/o acta circunstanciada:** en los casos en que se registra una averiguación previa o acta circunstanciada, el número de folio o número fue retirado de la Base Madre, y en la Base CIDE-PPD fue sustituido por una variable dicotómica que indica la existencia o ausencia de una AP o AC asociada con el evento. Mientras en las bases de **Enfrentamientos** y **Agresiones** se distingue entre AP y AC, en **Ejecuciones**

se incluyen las dos en la misma variable, por lo que la variable dicotómica indica si hubo una AP o AC asociada al evento.

Algunos de los cuadros externos de **Ejecuciones** son semejantes a los descritos para **Enfrentamientos** y **Agresiones**. Los cuadros externos creados exclusivamente para la categoría de **Ejecuciones** se encuentran descritos en la Cuadro 8 a continuación:

Cuadro 6: Cuadros externos de **Ejecuciones** (adicionales a los descritos en el Cuadro 6)

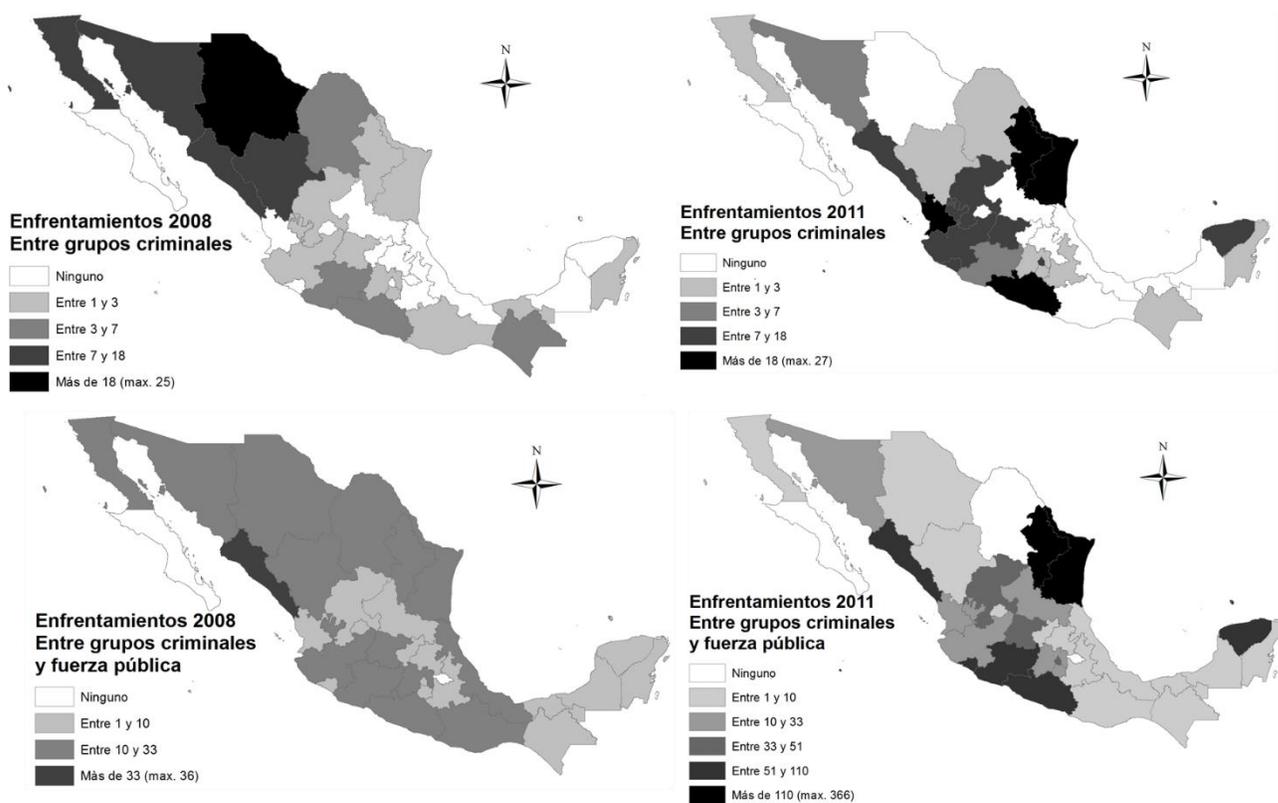
Número y nombre del cuadro	Información contenida
Tabla 3. Edad de los ejecutados	<p>Por lo general no existe información precisa en la Base Madre. Aunque algunos registros precisan la edad, en la mayor parte de los casos se trata de aproximaciones. Asimismo, cuando aparece más de una víctima por evento, no se registra la edad para todas las personas victimizadas. La mejor forma de extraer información útil de la Base Madre fue la creación de categorías de edad. Así, se incluyeron las siguientes:</p> <ul style="list-style-type: none"> • Menores de 15 años, • Entre 15 y 20, • Entre 20 y 30, • Entre 30 y 50, y • Mayores de 20 años de edad. <p>En cada categoría se indica simplemente si hubo víctimas en ese rango de edad, pero sin precisar cuántas víctimas pertenecen a cada uno de los rangos.</p>
Tabla 4. Tipo de arma(s) utilizada(s)	<p>En algunos eventos fue utilizada más de un arma para ejecutar a la víctima; por esta razón se creó una tabla externa para rescatar esta información, tomándola directamente de la Base madre y clasificando los tipos de armas en 15 categorías distintas.</p>
Tabla 5. Calibre-Munición	<p>Se identificaron 41 tipos distintos.</p>
Tabla 6. Modo de muerte	<p>La Base Madre tiene un campo que especifica el modo de muerte, pero la información allí contenida no es muy precisa. Por ejemplo, hay descripciones como “dentro de su domicilio” o “restos humanos”. Estas categorías no siempre aportan información valiosa, pero al ser la clasificación existente, se decidió incluirla. Se encontraron 30 modos distintos de muerte.</p>
Tabla 7. Clasificación de mensajes dejados al lado de los cuerpos	<p>El contenido de los mensajes que fueron dejados junto con los cuerpos fue uno de los espacios donde más información sensible se encontró, pues en algunos casos se refería a individuos específicos o se realizaban amenazas muy concretas. En consecuencia, la información de los mensajes se codificó siguiendo los criterios que propuso Atuesta (2015) a partir de la base original:</p> <ul style="list-style-type: none"> • Mensajes dirigidos al gobierno, • Mensajes dirigidos a informantes, • Mensajes que justifican los actos, • Mensajes de rivalidad entre grupos, • Mensajes relacionados con drogas o control territorial, • Mensajes de extorsión.
Tabla 8. Grupos involucrados en el evento	<p>Se identificaron 212 grupos criminales.</p>

Los diccionarios de variables para las tres bases de datos (al igual que sus tablas externas) se incluyen en el documento técnico que acompaña la base de datos, en el cual se precisa las diferentes categorías incluidas en cada una de las variables. La información obtenida de estas bases haría posible realizar un análisis geográfico y temporal sobre la evolución de la

violencia en México. Si bien el objetivo de este artículo no es efectuar dicho análisis, sino simplemente describir la Base CIDE-PPD, incluimos a continuación algunos mapas con un análisis temporal y espacial de algunas de las variables incluidas en la base de datos, a fin de ilustrar las aportaciones que ofrece la Base CIDE-PPD.

Los mapas incluidos en la Gráfica 8 muestran la evolución de los dos tipos de enfrentamientos mencionados anteriormente: aquellos en los que hubo involucramiento de alguna fuerza pública, y aquellos en los que se enfrentaron dos o más grupos criminales. Los mapas se realizaron con información de 2008 y 2011, respectivamente.

Gráfica 8: Evolución espacial y temporal (2008-2011) de los enfrentamientos entre grupos criminales y enfrentamientos entre grupos criminales y fuerza pública

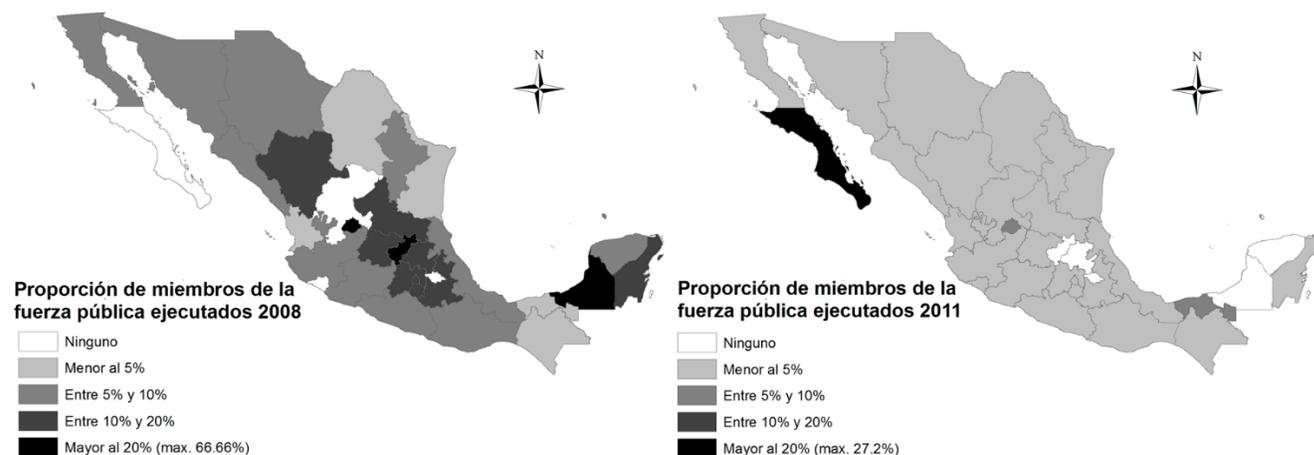


Como se observa en la Gráfica 8, las entidades con el mayor número de enfrentamientos entre grupos son también los estados que tuvieron el mayor número de enfrentamientos entre grupos criminales y fuerza pública. La evolución espacial de estos dos tipos de enfrentamientos es similar, pero también se observan algunas diferencias, principalmente en 2008: en ese año, los estados con el mayor número de enfrentamientos entre el gobierno y grupos criminales fueron Chihuahua, Tamaulipas, Sinaloa y Michoacán; mientras que los estados con el mayor número de enfrentamientos entre los grupos criminales fueron Chihuahua, Durango, Sinaloa, Sonora y Baja California. En 2011, Tamaulipas fue el estado que tuvo el mayor número de ambos tipos de enfrentamientos: entre la fuerza pública y los grupos criminales, y entre grupos criminales; seguido por Nuevo León, Guerrero y Nayarit.

Otro aspecto a considerar, al analizar la Gráfica 8, es el crecimiento en el número de enfrentamientos entre grupos criminales y el gobierno, comparado con el crecimiento en enfrentamientos entre grupos criminales. Mientras que el mayor número de enfrentamientos entre grupos criminales fue de 25 en 2008 (en Chihuahua), esa cifra se incrementó solo a 27 en 2011 (en Tamaulipas). En contraste, el mayor número de enfrentamientos entre grupos criminales y la fuerza pública fue de 36 en 2008 (en Sinaloa); mientras que en 2011, esta cifra se incrementó a 366 (en Tamaulipas), lo que representó un incremento de más de 900 por ciento.

La Gráfica 9 muestra el número de miembros de la fuerza pública ejecutados por estado en comparación con el número de civiles ejecutados para los años 2008 y 2011. Este análisis permite determinar quiénes son las principales víctimas de las ejecuciones, al igual que distinguir entre la violencia dirigida al gobierno y la violencia “privada” entre grupos criminales.

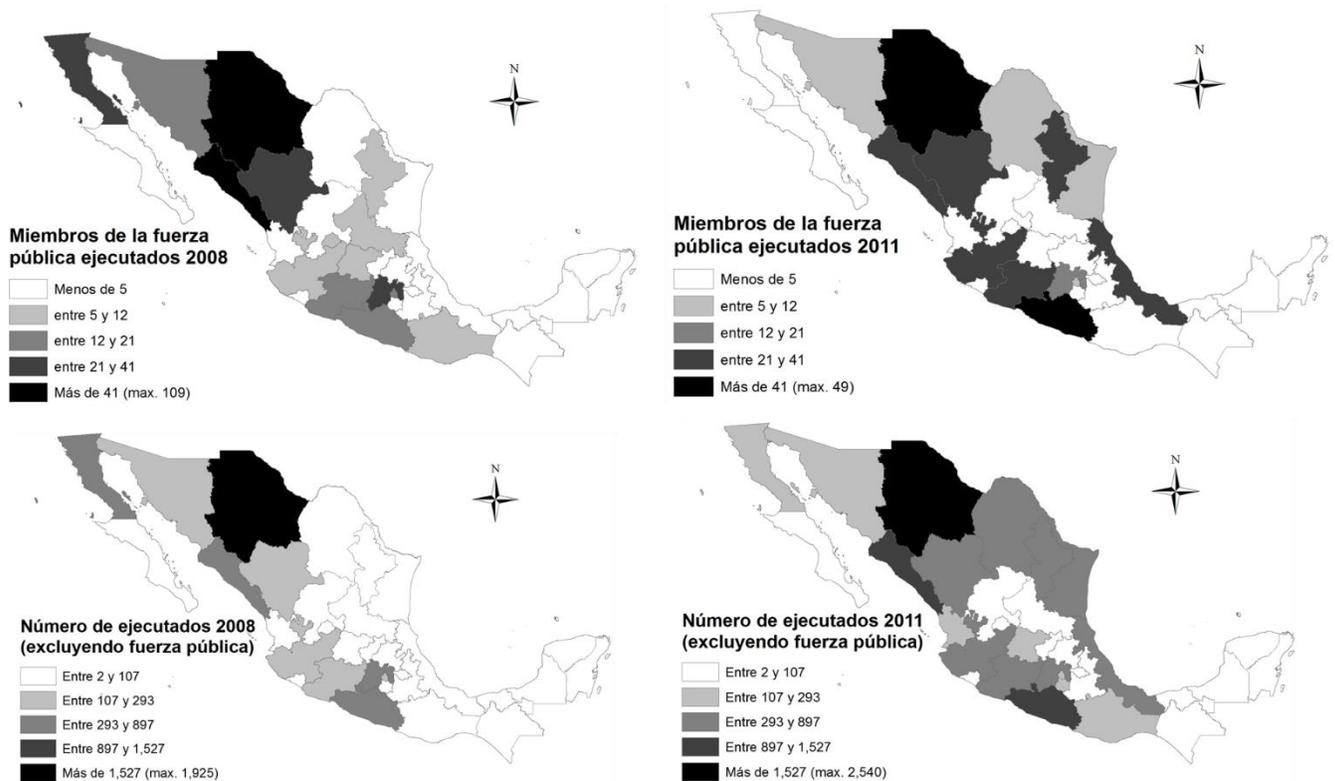
Gráfica 9: Proporción de miembros de la fuerza pública ejecutados por estado en 2008 y 2011



Las ejecuciones de miembros de la fuerza pública se incrementaron sustancialmente entre 2008 y 2011, pero ese crecimiento no fue homogéneo para todas las entidades federativas. En 2008, el mayor porcentaje de ese tipo de víctimas se observó en Querétaro (del total de seis ejecuciones, cuatro fueron miembros de la fuerza pública) y Campeche (tres de siete ejecuciones fueron de miembros de la fuerza pública); mientras que en 2011, el estado con la mayor proporción de víctimas de la fuerza pública sobre el total de ejecuciones fue Baja California Sur (de 11 ejecuciones, tres fueron miembros de la fuerza pública), seguido por Aguascalientes, el Distrito Federal, Michoacán, Tabasco y Veracruz (todos ellos con una proporción de fuerza pública ejecutada de 5 por ciento).

Tanto la proporción como el número total de miembros de la fuerza pública ejecutados disminuyeron de 2008 a 2011. La Gráfica 10 muestra esa evolución y la compara con la evolución del resto de ejecuciones.

Gráfica 10: Evolución espacial y temporal (2008-2011) de las **Ejecuciones** en las cuales la víctima es un miembro de la fuerza pública y aquellas en las que la víctima no es miembro de la fuerza pública



En términos absolutos, Chihuahua fue el estado con el mayor número de miembros de la fuerza pública ejecutados en 2008 (109) y el estado con el mayor número de ejecuciones (1,925). Mientras que el número de miembros de la fuerza pública ejecutados en ese estado se redujo a 47 en 2011, otras ejecuciones se incrementaron a 2,540 en el mismo periodo. Sinaloa, por su parte, fue el segundo estado con el mayor número de miembros de la fuerza pública ejecutados en 2008 (80), pero no en 2011, año en el que Guerrero y Chihuahua tuvieron las mayores víctimas (49 y 47, respectivamente, mientras que Sinaloa tuvo 25). Estos tres estados también tuvieron el mayor número de ejecuciones de civiles en 2011 (Chihuahua, 2,540; Guerrero, 1,517; y Sinaloa, 1,145).

Varias preguntas respecto a las consecuencias de la guerra contra las drogas pueden surgir al analizar estas gráficas. Por ejemplo, ¿cuáles fueron las implicaciones de la implementación de las operaciones militares con respecto a las víctimas de la violencia? ¿Son estas víctimas miembros de la fuerza pública, o civiles y presuntos delincuentes? ¿Cómo ha evolucionado la violencia en los diferentes estados? ¿Se puede hablar más del surgimiento de una violencia “privada” entre grupos criminales, o es esta violencia dirigida al estado? El objetivo de este artículo no es contestar estas preguntas o presentar recomendaciones de política. Nuestro interés reside solamente en presentar un análisis preliminar del tipo de investigaciones que se

pueden hacer con la información incluida en la Base CIDE-PPD, e invitar así a otros investigadores a explorar estas y otras preguntas, y sus posibles respuestas.

VI. Posibles sesgos y limitaciones de la Base de Datos CIDE-PPD

A pesar de que la Base CIDE-PPD incluye un gran cúmulo de información con respecto a la violencia generada por la guerra contra las drogas en México, es importante tener en cuenta que esta información contiene sesgos y limitaciones. Existen dos tipos de sesgos asociados con la Base CIDE-PPD: el primero está relacionado con la falta de fuente de la Base Madre. A pesar de que nosotros concluimos que la Base Madre corresponde a los registros administrativos utilizados para generar la Base Presidencia, estos archivos fueron filtrados al CIDE, por lo cual la Base Madre carece de fuente. El segundo sesgo está relacionado con la metodología utilizada para generar la información incluida en la Base Madre. Identificamos tres principales limitaciones relacionadas con dicha metodología: (i) los orígenes de la Base Madre; (ii) diferencias en cómo se recolectó la información dependiendo del cuerpo de la fuerza pública involucrado; (iii) diferencias espaciales y temporales en la información recolectada.

La razón principal por la cual el PPD realizó el ejercicio de validación fue para solucionar el primer sesgo. El objetivo de la validación era comprobar que la información contenida en la Base Madre no fue fabricada y que, efectivamente, correspondía a eventos que habían sucedido en el contexto de la llamada “guerra contra las drogas” en México. Con respecto al segundo sesgo, las tres limitantes mencionadas en el párrafo anterior se describen a continuación:

(i) Orígenes de la Base Madre

De acuerdo con la Guía metodológica publicada conjuntamente con la Base de Presidencia, la información fue recolectada por diferentes agencias federales desde 2009. La información fue revisada por dichas organizaciones y categorizada en función de los criterios establecidos para este objetivo. Debido a que las categorías creadas —**Enfrentamientos, Agresiones y Ejecuciones**— no tenían soporte legal, ya que no existen en el Código Penal Mexicano, el grupo encargado de realizar esa clasificación siguió un criterio y una metodología específicos para decidir qué homicidios debían ser incluidos en la Base Madre y cuáles deberían ser dejados afuera.

Cuando un evento era incluido en la Base Madre, se incluía también una descripción de los hechos. No está claro qué entidades o qué agencia del gobierno era la responsable de redactar esa descripción: pudieron haber sido la Sedena, las agencias que participaron en el enfrentamiento, las procuradurías locales, la Procuraduría General de la República, otros cuerpos de seguridad e inteligencia federales como el CISEN, o las autoridades que descubrieron los cuerpos de las víctimas ejecutadas. Sin embargo, después de analizar el contenido de las descripciones observamos que existe un posible sesgo en la forma en cómo se

describen los hechos. Por ejemplo, en los eventos incluidos en la base de datos de **Enfrentamientos** en los cuales hubo participación de algún cuerpo de la fuerza pública, *en casi la totalidad de los casos* la acción de ésta se describe como respuesta a un ataque de presuntos delincuentes o grupos armados. Es decir, según el reporte contenido en la Base Madre, el enfrentamiento sucedió como consecuencia de una agresión, pero nunca como una iniciativa de la fuerza pública. Un ejemplo de esta situación es la siguiente [descripción textual de un evento ocurrido en Tamaulipas en junio de 2011]:

“Elementos militares, al realizar patrullajes en el fraccionamiento (...), observaron un vehículo (...) y una camioneta (...), los ocupantes al ver la presencia militar emprendieron la huida, iniciando su persecución, siendo agredidos por disparos de armas de fuego de los ocupantes del vehículo (...), repeliendo la misma, mientras el vehículo (...) se dio a la fuga, al concluir la agresión procedieron a revisar el lugar encontrando abandonado el vehículo (...), el cual al ser revisado se encontró un arma larga, cintas metálicas de eslabones abiertos para cartuchos calibre .308 y cartuchos útiles”.

Varias preguntas surgen de esta descripción: (i) ¿Cómo detectó el personal militar a estos individuos?; (ii) ¿Si los individuos armados empezaron la agresión, por qué no hubo militares heridos?; (iii) ¿Los militares trataron de interactuar con los individuos antes de que se abriera el fuego? Sin querer insinuar que los eventos sucedieron de una manera distinta, es improbable que en ninguno de los eventos los cuerpos de la fuerza pública involucrada dieran inicio al enfrentamiento, y que éstos sucedieran siempre porque presuntos delincuentes atacaron primero a miembros de la fuerza pública.

Otro ejemplo se incluye a continuación [descripción textual de un evento sucedido en Aguascalientes, en junio de 2007]:

“En Aguascalientes, tripulantes de una camioneta pick up 4x4 que estaba a punto de ser revisada en un retén policial se dieron a la fuga desatando una persecución que acabó en tiroteo.”

De la descripción de este evento, la Base Madre identifica a los tripulantes de la camioneta como narcotraficantes, pero no se reporta ningún detenido, herido, o muerto en la agresión; no se reporta que se hubiera registrado o asegurado ningún cargamento; y no se provee de mayor información que permitiera corroborar la clasificación de “narcotraficantes” de los tripulantes de la camioneta, por lo que no se puede saber exactamente qué fue lo que sucedió. En este caso, la fuerza pública involucrada fue el Ejército, que empezó la persecución pero no detuvo a nadie. Si esto fue así, ¿con base en qué se confirmó que los integrantes de la camioneta eran narcotraficantes? ¿Encontraron algún cargamento de drogas pero no lo reportaron en la Base Madre? Si esto fue así, ¿qué información, y bajo qué criterios, se incluyó el evento en la Base Madre?

Con respecto a la recolección de la información para ser incluida en la Base Madre, la Guía metodológica explica que “[c]ada integrante del grupo aporta información fidedigna y confiable de sus despliegues estatales que se encuentra relacionada con muertes y eventos de enfrentamiento o agresión que ocurrieron, presumiblemente, en el contexto de la rivalidad

entre grupos delincuenciales” (p. 3). Nosotros solo sabemos que hay dependencias gubernamentales responsables de recolectar la información y rellenar el campo de descripción de los hechos. Por esta razón no es sorprendente que, en la descripción de los eventos, la fuerza pública involucrada siempre estuviera “repeliendo la agresión” o defendiéndose de los presuntos delincuentes.

- (ii) Diferencias en la información recolectada dependiendo de la fuerza pública involucrada

A pesar de que ya sabemos que los encargados de recolectar la información para la Base Madre eran órganos encargados de la fuerza pública o de la inteligencia gubernamental, éstos siguieron un criterio de selección de los eventos que, si bien es detallado en la Guía, parece más indicativo que discreto. En consecuencia, no podemos afirmar que todas las agencias hayan usado el mismo criterio para recolectar la información. En otras palabras, la descripción de los eventos (el grado de detalle, la longitud y la calidad de esta información) depende de la agencia responsable de hacerlo; probablemente los criterios de clasificación también varíen dependiendo de la agencia responsable de capturar la información.

No siempre se sabe qué agencia gubernamental es la responsable de recolectar esta información. Podemos suponer que el cuerpo de la fuerza pública que estuvo involucrado en el evento es el responsable de recolectar la información, pero no hay forma de verificarlo. Por lo pronto, en los casos de **Ejecuciones** o **Enfrentamientos** entre grupos criminales, donde no hubo participación de la fuerza pública, por lo general no hay indicios sobre quién fue el responsable de recolectar la información. Usando la base de datos de **Enfrentamientos** entre la fuerza pública y grupos criminales, podemos darnos una idea de las variaciones en la forma en que la información recolectada es capturada dependiendo de la agencia involucrada. Algunas de las características sobre esta recolección de información son resumidas en el Cuadro 9 continuación:

Cuadro 7: Características del campo de descripciones en **Enfrentamientos** cuando diferentes cuerpos de la fuerza pública estuvieron involucrados

Cuerpo de la fuerza pública involucrado en el Enfrentamiento	Características sobre la descripción de los eventos
AFI	<p>En la mayoría de los casos se utiliza la frase “repelieron la agresión”. Un evento se describe como un grupo de enfrentamientos en donde participaron agentes de la fuerza pública. Por ejemplo, un enfrentamiento que se reportó después de que la AFI aseguró un área en donde se encontraron vehículos abandonados (Morelos; mayo, 2008).</p> <p>También se “repelen agresiones” cuando se llevan a cabo actividades especiales, tales como garantizar la protección a testigos (Durango; mayo, 2009), o cuando tratan de rescatar individuos secuestrados (Estado de México; marzo, 2009).</p>
Semar	<p>En la mayoría de los casos, los marinos “repelen una agresión” cuando ven a individuos armados en un vehículo, cuando están llevando a cabo actividades</p>

	<p>de vigilancia aérea o terrestre, o cuando presuntos criminales huyen al ver llegar a los marinos.</p> <p>En otros casos, los marinos llegan al sitio de los eventos cuando dos grupos criminales están enfrentándose y, al verlos llegar, los presuntos criminales huyen del sitio.</p>
Sedena	<p>En la descripción de los hechos normalmente se utiliza la frase “personal militar sin novedad” cuando no hubo militares víctimas (muertos o heridos).</p> <p>Aproximación de vehículos a alta velocidad: en la mayoría de las descripciones analizadas los militares no atacan primero, sino que “repelen la agresión” o son “objeto de una agresión”. Sin embargo, no siempre la descripción de la “agresión” es incluida. Una “agresión” puede ser un vehículo que se está aproximando a alta velocidad (por ejemplo, “personal militar que atendía una denuncia anónima, al transitar por la carretera (...), se encontró con un bloqueo vehicular (<i>trailers</i>), por lo que se adoptó un dispositivo de seguridad, detectando un vehículo particular que se aproximaba a alta velocidad, <i>por lo que se repelió la agresión</i>. Derivado de los hechos se registró la muerte de un civil, quien vestía una camisa y una chamarra con logotipos de (...). Tamaulipas; febrero, 2010).</p> <p>En otras descripciones, los militares abren fuego en contra de individuos que mostraban una conducta sospechosa: “En la zona de (...) frente al (...), a escasos (...) un enfrentamiento a balazos entre elementos del Ejército Mexicano y dos civiles, tripulantes de un auto (...), quienes intentaron darse a la fuga luego de que los militares intentaron detener el auto (...) para hacerle una revisión dado que los dos hombres que lo tripulaban mostraban actitudes sospechosas. Uno de los delincuentes sacó de entre sus ropas una pistola, <i>lo que fue tomado como una actitud de agresión por los militares</i> que de inmediato accionaron sus armas en contra de los ocupantes del vehículo, logrando detener a los dos individuos” (Veracruz; abril, 2008).</p> <p>Una persecución también es considerada una agresión, por lo cual se justifica abrir fuego contra los presuntos delincuentes: “elementos de la Secretaría de la Defensa Nacional detectaron en las inmediaciones del poblado de (...), en (...) al menos dos camionetas en ‘convoy’ con sujetos armados, mismos que al percatarse de la presencia de los soldados emprendieron la huida, dando inicio a una persecución. <i>Por lo que fue repelido el ataque</i>, finalmente dos sujetos escaparon entre el monte y los dos restantes fueron capturados (Sinaloa; julio, 2010).</p>
Policía Estatal	<p>En algunos casos se “repele una agresión” pero en la mayoría, la descripción no especifica cómo empezó el enfrentamiento. Se describe el evento como “se registró un enfrentamiento”, “la Policía estatal se enfrentó contra...”, o “protagonizaron un enfrentamiento”.</p>
Policía Federal	<p>De los eventos analizados, en el caso de la Policía Federal no pudimos encontrar ninguno en el cual se estuviera “repeliendo una agresión”. La descripción solo dice que “se llevó a cabo un enfrentamiento”, “se registró un enfrentamiento”, “se enfrentaron”, o “hubo un enfrentamiento como resultado de una orden de aprehensión”.</p>
Policía Federal Ministerial	<p>Solo se registra el enfrentamiento usando palabras tales como “registró”, “enfrentó”, o “como resultado de una orden de aprehensión”.</p>

	<p>Cuando hubo una agresión antes del enfrentamiento, usualmente la descripción incluye alguna información sobre la agresión: ya sea que la agresión fue en contra de un agente específico o de un edificio del gobierno.</p>
Policía Ministerial	<p>Los eventos se describen “como parte de un enfrentamiento”; “se registró un enfrentamiento”, o “repelieron una agresión”. En algunos casos fueron emboscados o recibieron una falsa alarma; en otros, empezaron una persecución después de una agresión y luego tuvo lugar el enfrentamiento.</p> <p>En un evento específico se menciona un enfrentamiento en contra de los militares: “Supuestos militares quienes se encontraban en su día de descanso ejecutaron a un civil y desataron un enfrentamiento en contra de Ministeriales, matando a tres agentes. Los supuestos soldados levantaron a dos hombres en una pista para avionetas a lampazos, en donde ejecutaron a uno de los secuestrados y se robaron una camioneta en la que huyeron. Se les descompuso el vehículo por lo que robaron otro automóvil para continuar su huida, en (...) se topan con ministeriales, a quienes rafaguearon, muriendo dos... (Nuevo León; abril, 2008).</p>
Policía Municipal	<p>La descripción es muy completa. Cuando “repelen una agresión”, se incluye la información acerca de la agresión. En algunas ocasiones responden a una llamada anónima o a un reporte, o están tratando de liberar a individuos secuestrados.</p> <p>En algunos casos la descripción de los eventos parece estar escrita por una tercera parte. Por ejemplo: “a las 04:45 horas por medio de una llamada telefónica al número de emergencias 066 con número de folio (...), se tuvo conocimiento que elementos de la Policía Municipal, mientras realizaban un patrullaje de rutina a bordo de la unidad (...), en el cruce (...), sujetos que viajaban a bordo de una camioneta (...) los agredieron con disparos de arma de fuego, repeliendo la misma y emprendiéndose una persecución, dándoles alcance en (...), dando como resultado tres personas aseguradas (...)” (Chihuahua; noviembre, 2011).</p> <p>De este caso específico no se sabe a ciencia cierta cómo una llamada anónima pudo haber sido tan específica en cuanto a los hechos del enfrentamiento (dado que se describen exactamente las ubicaciones de la agresión y del posterior enfrentamiento).</p> <p>En otros casos, el enfrentamiento tuvo lugar después de un enfrentamiento entre dos grupos armados (por ejemplo; Guanajuato, junio, 2010).</p>

A pesar de que hay diferencias en el campo de descripción de los hechos dependiendo del cuerpo de la fuerza pública involucrado en el evento, en la mayoría de los casos se refieren al enfrentamiento como una consecuencia de una agresión previa y lo describen como “repeliendo la agresión”. La agresión puede ser por medio de armas de fuego, una persecución, un vehículo aproximándose muy rápidamente o un individuo o vehículo sospechoso. En algunos casos (sobre todo cuando hay participación de la Semar) se describe a los individuos como “civiles”, pero en otras ocasiones (particularmente cuando hay participación de la Sedena) se describe a los individuos como criminales, inclusive cuando se presenta un contexto familiar: hombres viajando con, o acompañados de, sus esposas e hijos.

Cuando hay participación de las fuerzas policiales (Federal, Ministerial, Estatal o Municipal) no se incluye información sobre el inicio del enfrentamiento. Solo en algunos casos se “repele la agresión”, pero usualmente la descripción incluye frases como “se registró un enfrentamiento”, “se enfrentaron”, o “protagonizaron un enfrentamiento”. De todas maneras, no es claro si hubo un criterio para la forma en cómo el campo de descripción de los hechos fue escrito, por quién y cuándo. En algunos casos parece que se escribió apenas tuvo lugar el evento; en otros parece haber sido escrito después de que se recibió una llamada anónima que ofreció los detalles del enfrentamiento; y en otros casos, parece haber sido escrito por una tercera persona que no tuvo participación en el evento.

A pesar de que el campo de descripción fue eliminado de la Base de Datos CIDE-PPD (dado que contiene información sensible), es importante entender cómo fue registrado el evento y bajo qué enfoque. Si en apariencia los cuerpos de la fuerza pública *siempre* son atacados antes de un enfrentamiento, la mayoría de los homicidios registrados en estos eventos serán registrados como fallecimientos de presuntos criminales y no de civiles, a pesar de que no quede claro cuál fue la agresión en primer lugar.

(iii) Diferencias espaciales y temporales en la información recolectada

Las diferencias en cómo los eventos se describen y la información recolectada no depende solamente del cuerpo de la fuerza pública involucrado, sino también del lugar y el momento en el cual el evento tuvo lugar. Una de las principales diferencias de la información recolectada a través del tiempo es el número de variables incluidas en la Base Madre. En el caso de los **Enfrentamientos** y las **Agresiones**, las variables son las mismas de 2006 a 2011, pero en el caso de las **Ejecuciones**, nuevas variables fueron incluidas a través del tiempo y después removidas de la Base Madre. El Cuadro 10 resume estos cambios observados en las variables incluidas en la base de datos de **Ejecuciones**.

Cuadro 8: Variables incluidas en los diferentes años en la categoría de **Ejecuciones**

Ejecuciones 2006	Información no incluida: <ul style="list-style-type: none"> • Edad años • Cargo • Policías y/o funcionarios públicos • Información sobre los grupos criminales involucrados en el evento (perteneciente a la organización delictiva y organización delictiva que cometió el delito) • Modo de muerte • Averiguación previa (y/o acta circunstanciada) • Observaciones (adicionales a las características de los hechos)
Ejecuciones 2007	Información no incluida: <ul style="list-style-type: none"> • Edad años • Cargo • Policías y/o funcionarios públicos • Información sobre los grupos criminales involucrados en el evento (perteneciente a la organización delictiva y organización delictiva que cometió el delito) • Modo de muerte • Averiguación previa (y/o acta circunstanciada)

	<ul style="list-style-type: none"> • Observaciones (adicional a las características de los hechos)
Ejecuciones 2008	Se incluyeron todas las variables
Ejecuciones 2009	Se incluyeron todas las variables
Ejecuciones 2010	Se incluyeron todas las variables
Ejecuciones 2011	Información no incluida: <ul style="list-style-type: none"> • Observaciones (adicionales a las características de los hechos)

Las principales diferencias se observan en los primeros dos años, 2006 y 2007, cuando algunas variables sobre las características de las víctimas, información de los grupos participantes y características de los eventos no fueron recolectadas. Esta falta de información representa una limitante para la Base de Datos CIDE-PPD en el momento en que requiera para hacer un análisis temporal, porque la información no recolectada puede ser interpretada como si no existiera. Por ejemplo, cuando se analiza el número de policías y/o funcionarios públicos ejecutados, la única información disponible es si hubo o no policías ejecutados, pero no cuántos fueron (en 2006 y 2007). O cuando se analizan los grupos involucrados en cada evento, si el nombre del grupo no está incluido en las características de los hechos, o en los mensajes que fueron dejados al lado de los cuerpos, no se sabe si no hubo grupos involucrados o si simplemente no fueron identificados.

Analizando las características de los eventos, nos percatamos de que la narrativa también cambia a través del tiempo, con más detalles en algunos años que en otros. En diciembre de 2006 (primer año con información), el promedio del número de palabras en las características de los hechos (en la base de datos de **Ejecuciones**) fue de 19. Este número se incrementó a 33.3 en 2007, a 35.9 en 2008, a 37.67 en 2009, a 43.07 en 2010, y finalmente a 45.15 en 2011. La mayor tasa de crecimiento en el número de palabras incluidas en las características de los hechos fue de 2006 a 2007, un incremento de 75 por ciento. A pesar de que la longitud de la descripción no está necesariamente asociada con una mejor descripción de los hechos, en la mayoría de los casos, este supuesto sí se cumple. El Cuadro 11 a continuación muestra diferentes ejemplos de las características de cada año, para observar estos cambios en la narrativa.

Cuadro 9: Características de los hechos en la base de datos de **Ejecuciones** seleccionadas aleatoriamente (2006-2011)

Diciembre 2006 - Chihuahua	<i>“Fue ejecutado por sujetos armados en la clínica (...).”</i>
Octubre 2007 - Guerrero	<i>“El director de la Policía Investigadora Ministerial de Guerrero (...), reportó que (..) fue asesinado de dos tiros en (...).”</i>
Mayo 2008 - Chihuahua	<i>“La víctima viajaba en un auto (...), cuando sicarios lo obligaron a detenerse chocándolo. Al detener la marcha, los victimarios abrieron ambas portezuelas y le dispararon al menos 33 veces”.</i>
Septiembre 2009 - Michoacán	<i>“Fue asesinado un hombre de 33 años de edad, custodio del Centro de Readaptación Social, cuando se encontraba en el exterior de su domicilio acompañado de su esposa, quien resultó ilesa. El ataque se registró en el</i>

	<i>momento que un grupo armado intentó ‘levantarlo’”.</i>
Enero 2010 - Sinaloa	<i>“Personas armadas a bordo de una camioneta (...) privaron de la vida a la víctima, quien quedó a bordo de un automóvil (...), y presenta 12 disparos de arma de fuego en diferentes partes del cuerpo. En el lugar se recogieron 26 casquillos percutidos calibre 9 mm y 4 ojivas del mismo calibre.”</i>
Enero 2011 - Sonora	<i>“Dos personas viajaban a bordo de una camioneta pick up (...), quienes eran seguidos por un grupo armado que viajaba en dos vehículos: (...), fueron asesinadas al llegar a la confluencia de (...). Los occisos presentaron impactos de arma de fuego en distintas partes del cuerpo. En el lugar se localizaron casquillos percutidos de armas de fuego tipo R-15, AK-47, un casquillo de 38 súper y uno de .9 milímetros.”</i>

También se observan diferencias cuando se analizan la características de los hechos a través del tiempo. Por ejemplo, en la descripción de los hechos presentada en diciembre de 2006 en Chihuahua no se identifica el sexo de la víctima, el modo de muerte o los motivos. En 2007, a pesar de que se registra el modo de muerte (y el nombre de la víctima), no se incluyen más datos acerca de la descripción del sitio o información acerca de los victimarios. En 2009 se sabe algo de la víctima (su cargo y su edad), al igual que el lugar donde estaba cuando fue asesinado. También se incluye alguna información acerca de los victimarios: miembros de un grupo criminal que buscaban secuestrarlo. En 2010 se describe cómo se encontró la víctima y los casquillos que se hallaron en el lugar. Una descripción similar se presenta en 2011, que incluye información más exacta sobre los vehículos, el lugar exacto de los hechos y la situación en la cual se produjo el asesinato.

Como se comentó anteriormente, estas características de los hechos no se incluyen en la Base CIDE-PPD, pero de ese campo se obtuvieron nuevas variables, tales como si hubo fuego cruzado, levantamientos, llamadas, drogas, vehículos, y grupos involucrados, entre otras (ver sección X). Por esta razón, es posible que la información esté más completa en algunos años que en otros dependiendo de la calidad de la información que se haya recolectado en el campo de características de los hechos, lo que sesgaría de esta manera el análisis temporal.

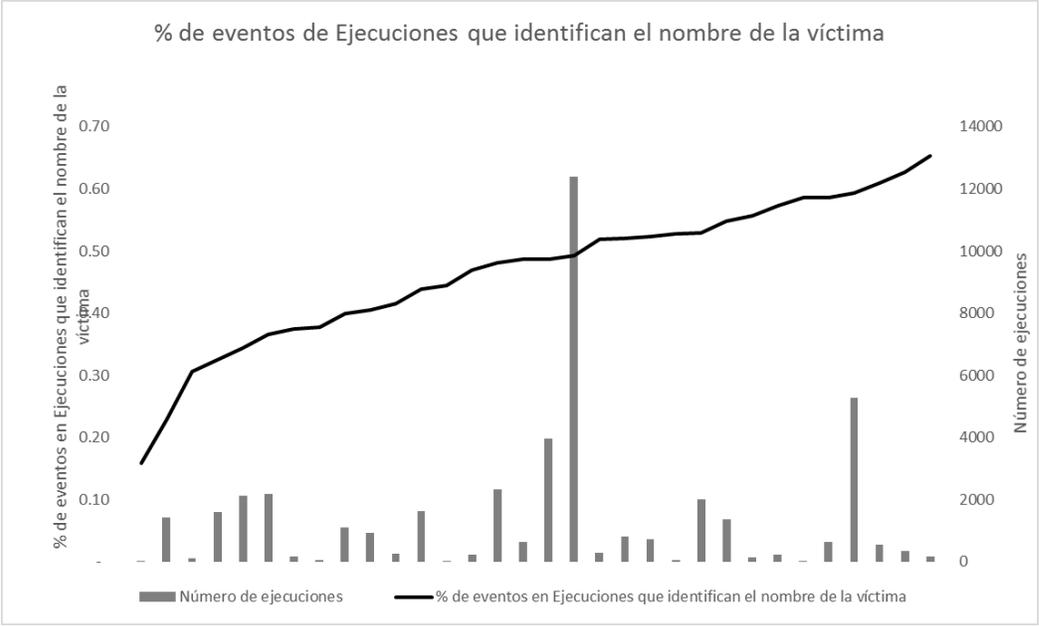
También se observa información heterogénea cuando se hace una comparación estatal. De acuerdo con la Guía metodológica, los despliegues estatales de cada integrante del grupo encargado de construir la Base Madre aportaban información fidedigna y confiable, y esta información se clasificaba en **Agresiones, Enfrentamientos o Ejecuciones**. La información también podía ser excluida de la Base Madre si una investigación posterior, responsabilidad del Ministerio Público, determinaba que se trataba de un evento con características distintas a la categorización hecha por el grupo. La narración de las descripciones para los eventos que sucedieron en cada uno de los estados es diferente, y en algunos casos es más específica que en otros. Como el 85 por ciento de los eventos corresponde a **Ejecuciones**, el análisis temporal se hará solamente para esta base de datos.

Hay diferentes formas en las cuales se puede analizar esta limitación. Primero, podemos contar el número de eventos en los cuales se incluye el nombre de la víctima. Si el nombre de

la víctima es conocido, se puede recolectar más información del individuo. Segundo, se puede contar el número promedio de palabras que se usan en las características de los hechos en cada estado para determinar qué estados describen más detalladamente los eventos. Finalmente, se puede contar el número de eventos en los cuales se incluye información adicional (en el campo de observaciones). A pesar de que este campo no está disponible todos los años, en los años en los cuales sí se incluye (de 2008 a 2010), no siempre se incluye información y se pueden observar las diferencias entre los estados.

La Gráfica 11 muestra el número total de **Ejecuciones** de 2007 a 2011 por estado y la proporción de los eventos en los cuales se incluyó información con respecto al nombre de la víctima.

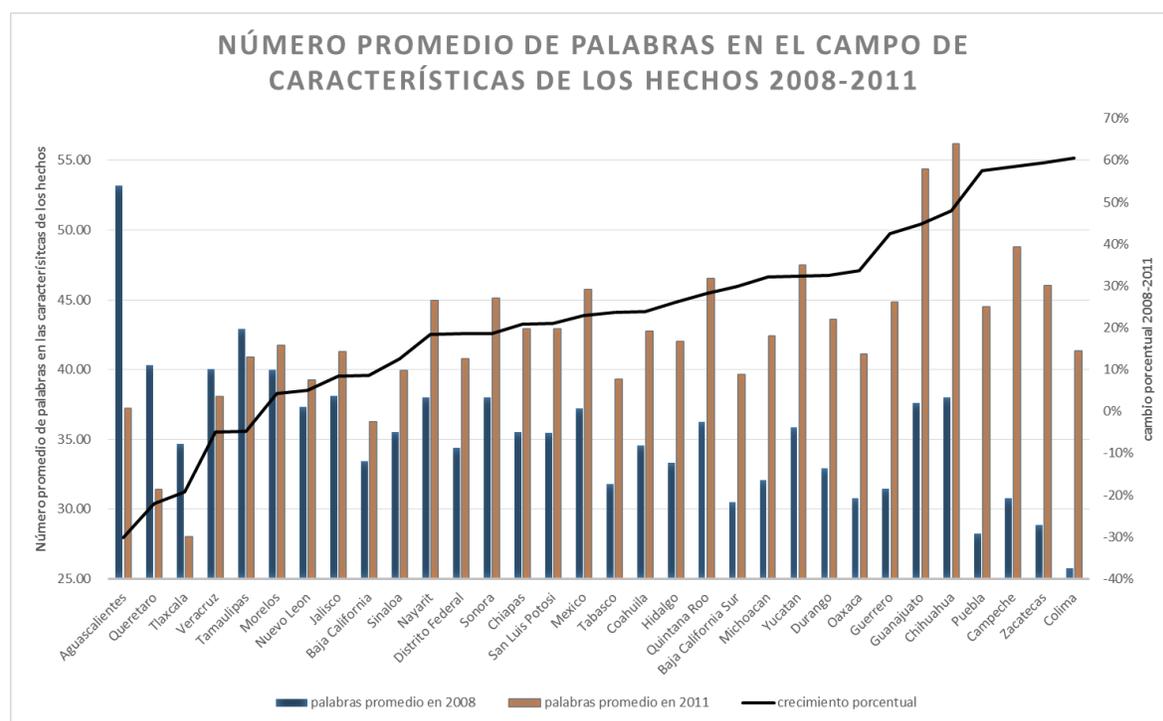
Gráfica 11: Número total de **Ejecuciones** de 2007 a 2011 por estado y proporción de los eventos en los cuales se incluyó información con respecto al nombre de la víctima



prensa en cada estado, la ocupación de las víctimas (víctimas más conocidas serían más identificables que otras), o qué tan peligroso pudiera ser para la fuerza pública reportar esta información si los grupos del crimen organizado no quisieran que fuera pública.

Para el análisis sobre el número de palabras incluidas en el campo de características de los hechos, usamos los eventos de 2008 y 2011 para realizar la comparación. Como se dijo anteriormente, el promedio nacional del número de palabras incluido en este campo en 2008 era de 35.9, mientras que en 2011 se incrementó a 45.15. La información por estado, tanto en 2008 como en 2011, se observa en la Gráfica 12.

Gráfica 12: Número promedio de palabras en el campo de características de los hechos por estado y su cambio porcentual de 2008 a 2011



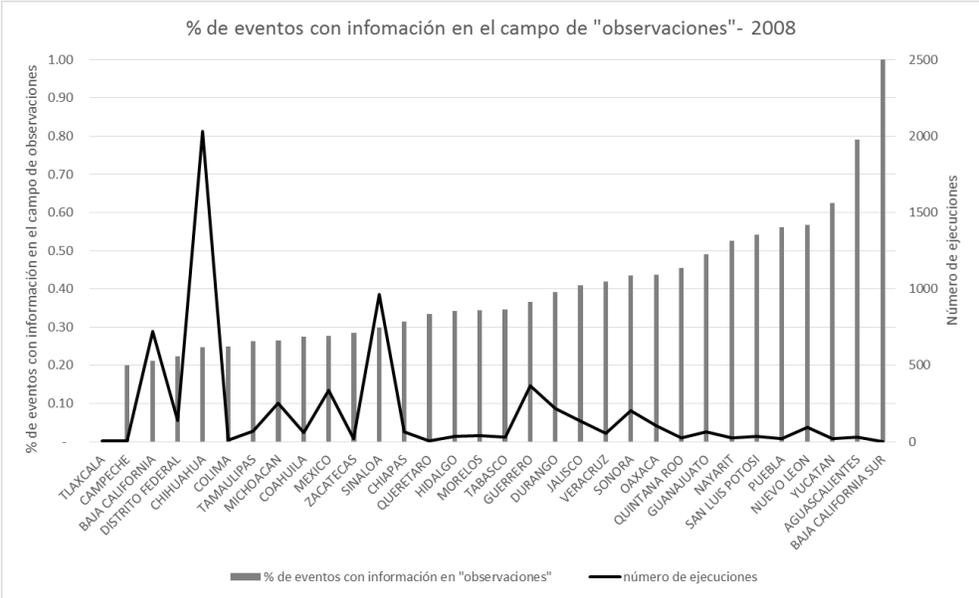
Finalmente, usando la categoría de **Ejecuciones** para los años 2008, 2009 y 2010, se realizó un conteo de los eventos en los cuales se incluyó información en el campo de “observaciones”. Usualmente en este campo se incluye información adicional de testigos, o información previa sobre la víctima. Ejemplos de las narrativas que se incluyen en este campo se presentan a continuación:

Sinaloa, enero de 2009: “La víctima fue levantada el (...), por un grupo armado que irrumpió en su domicilio en la comunidad de (...).”

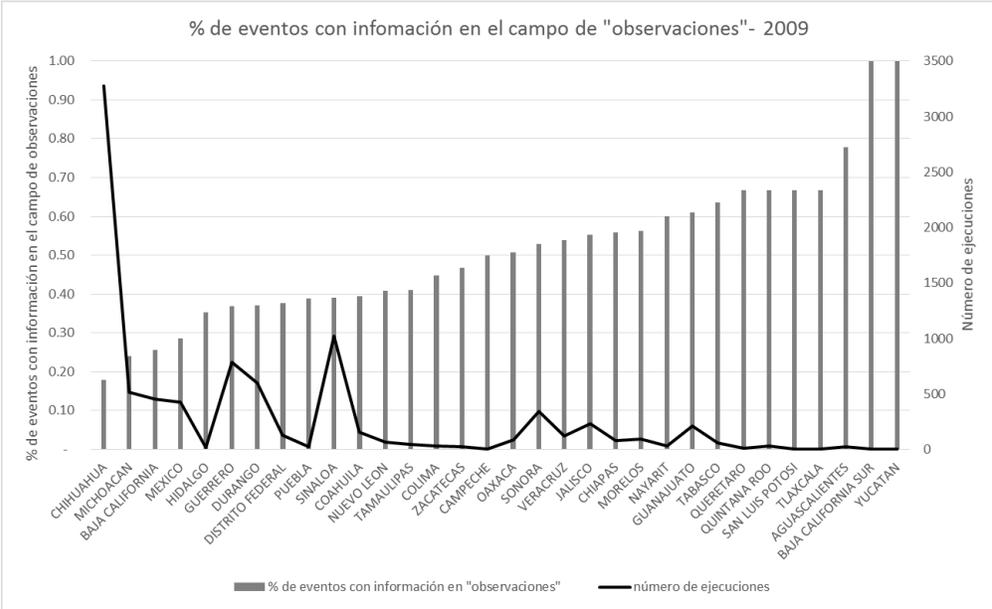
Veracruz, febrero de 2010: “Las autoridades ministeriales no descartan que estos hechos [se] haya[n] tratado de un ajuste de cuentas por parte del crimen organizado, de la célula de los Zetas que trabaja para el Cártel del Golfo y que tiene presencia en el sur del estado”.

Para los años 2008, 2009 y 2010, los porcentajes de eventos en los cuales se incluyó información en el campo de observaciones a nivel estatal se presentan en las gráficas 13, 14 y 15 a continuación:

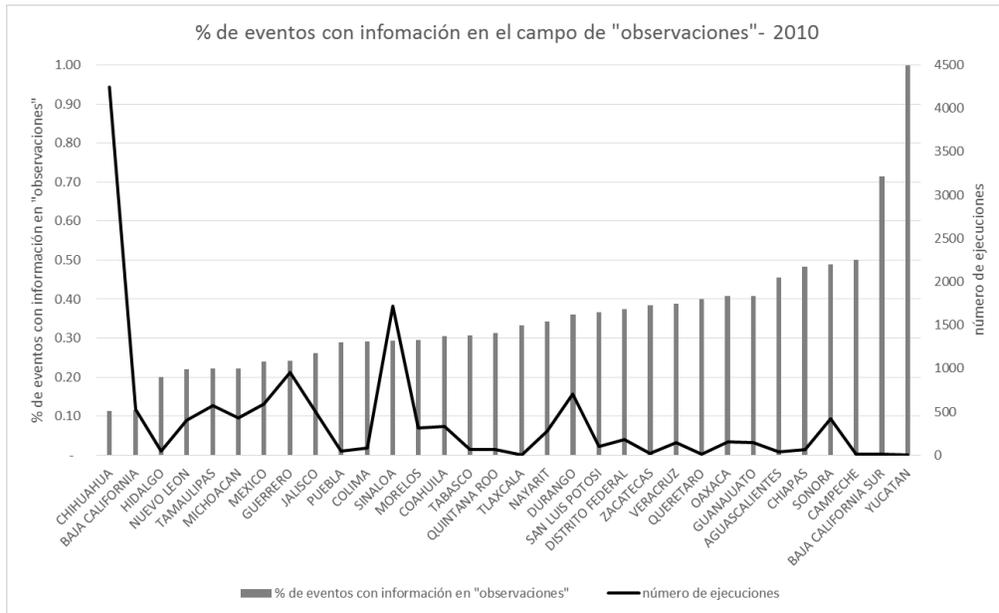
Gráfica 13: Porcentaje de eventos con información en el campo de “observaciones” en 2008 y número de ejecuciones por estado



Gráfica 14: Porcentaje de eventos con información en el campo de “observaciones” en 2009 y número de ejecuciones por estado



Gráfica 15: Porcentaje de eventos con información en el campo de “observaciones” en 2009 y número de ejecuciones por estado



De 2009 a 2010, los estados con los porcentajes más bajos de eventos con información en el campo de observaciones fueron Chihuahua y Baja California (incluyendo también a Michoacán en 2009), y éstos fueron también los estados con el mayor número de ejecuciones. A pesar de que no existe una relación directa entre estas dos variables (información en el campo de observaciones y número de ejecuciones), se puede observar alguna relación: los estados en donde hubo más **Ejecuciones** fueron los menos propensos a incluir información adicional en el campo de observaciones. La excepción a esta tendencia se observa en Sinaloa, estado que presenta un alto número tanto de ejecuciones como de eventos en los cuales se incluye información en el campo de “observaciones”: en 39 por ciento y 29 por ciento de los eventos en 2009 y 2010, respectivamente, se incluyó información adicional en “observaciones”.

Por otra parte, los estados con el menor número de ejecuciones (Campeche, Baja California Sur, Yucatán y Aguascalientes) tienen el mayor número de eventos con información en el campo de “observaciones”. Contrario a lo que sucede con los análisis que se realizaron previamente (identificación del nombre de las víctimas y el número de palabras en el campo de características de los hechos), este análisis sugiere que se provee más información sobre la víctima cuando menos ejecuciones se observan en el estado.

En este apartado se describieron las principales limitaciones de la Base Madre, y consecuentemente de la Base de Datos CIDE-PPD. Sin saber exactamente qué mecanismos y qué criterios fueron usados para recolectar la información, no es posible determinar cuál procedimiento es mejor para solucionar el problema de información incompleta. Little y Rubin (2002) describen diferentes opciones para resolver problemas de información incompleta o *missing data* que pueden tenerse en cuenta para disminuir el sesgo generado por la falta de información. Es importante que los investigadores que usen la Base de Datos CIDE-PPD

tomen en consideración estas limitaciones al realizar sus análisis y traten de presentar alternativas para solucionarlos. Por ejemplo, tratar de cruzar la información públicamente disponible con la información en la Base de Datos CIDE-PPD para solucionar los problemas de información incompleta, o utilizar técnicas estadísticas para solucionar el sesgo generado por la falta de información.

VII. Conclusiones

En este artículo se describe el proceso de validación y codificación de la Base Madre para la creación de una nueva base de datos denominada Base CIDE-PPD. En esta nueva base se incluye información sobre los eventos violentos relacionados con la guerra contra las drogas durante el gobierno de Felipe Calderón. Adicionalmente se provee de información con la que se puede vincular la Base Madre con la Base de datos de fallecimientos ocurridos por presunta rivalidad delincinencial, publicada por la Oficina de Presidencia en enero de 2011.

La información incluida en la Base CIDE-PPD permite realizar análisis de violencia relacionada con la guerra contra las drogas y probar hipótesis vinculadas con el tema que no se han podido analizar con la información hasta el momento disponible. Ejemplos de posibles análisis que pueden ser desarrollados con esta nueva base de datos se muestran en los apartados anteriores. Sin embargo, es importante que aquellos que decidan usar esta base de datos para sus análisis de violencia en México tengan en cuenta sus sesgos y limitaciones al momento de sacar conclusiones.

La codificación de la Base Madre todavía se encuentra en proceso, para poder extraer más información de los campos sensibles que incluye. En un futuro se realizarán actualizaciones de la Base CIDE-PPD, que estarán disponibles para uso público. Tanto el proceso de validación, como el de codificación de la Base Madre se realizaron para proteger a los individuos involucrados en la guerra contra las drogas y evitar así un uso irresponsable de los datos publicados.

[Actualizadas conforme a los lineamientos 2015 del CIDE en archivo aparte]

Anexo 1. Discrepancias entre la Base Madre y la Base CIDE-PPD

	Agresiones			Enfrentamientos			Ejecuciones		
	Mes	Base Presidencia	Base Madre	Mes	Base Presidencia	Base Madre	Mes	Base Presidencia	Base Madre
2006									
2007							1	135	133
							2	85	84
							3	223	222
							6	193	194
							9	239	238
							10	267	270
2008							12	250	253
							1	276	275

							2	301	302
							3	308	309
							7	557	560
2009				8	50	49	1	558	557
				9	57	58	3	576	575
							4	572	571
							8	818	822
							9	868	870
							10	834	837
							11	794	797
							12	973	975
2010	6	31	34	7	200	197	2	849	851
				12	98	99	3	1,039	1,038
							5	1,248	1,245
							7	1,263	1,260
							8	1,275	1,274
							9	980	977
							12	1,023	1,015
2011	1	No hay información	17	1	No hay información	208	1	No hay información	1,139
	2		15	2		197	2		957
	3		16	3		257	3		1,149
	4		31	4		216	4		1,381
	5		27	5		287	5		1,222
	6		21	6		147	6		1,259
	7		42	7		287	7		1,191
	8		73	8		207	8		1,180
	9		40	9		150	9		1,199
	10		36	10		208	10		1,263
	11		57	11		120	11		982