GVA Project Onboarding: Police Firearm Injury Data

February 25, 2026

Contents

1 Overview

This document explains the data source, case selection criteria, and cleaning and coding process used in this project's research on firearm assaults on U.S. police officers. It is intended for new research assistants, collaborators, and others who will work with or need to understand the project's data.

The core of this project relies on data collected by the **Gun Violence Archive (GVA)**, which we manually clean and recode to produce an analytic dataset of fatal and nonfatal firearm assaults on U.S. police officers in the line of duty.

---

2 The Gun Violence Archive (GVA)

2.1 What is GVA?

GVA is an independent, nonpartisan nonprofit organization whose mission is to "provide free online public access to accurate information about gun-related violence in the United States" (GVA, cited in Sierra-Arévalo

& Nix, 2020, p. 1044). GVA's definition of gun-related violence is broad, encompassing firearm homicides, suicides, injuries, accidental shootings, defensive firearm uses, mass shootings, officer-involved shootings, and more.

To gather its data, GVA staff use a mixture of automated and manual techniques to monitor more than 7,500 sources, including local and state law enforcement agencies, social media platforms (e.g., Twitter and Facebook), local and national news agencies, and governmental sources (Sierra-Arévalo et al., 2023, p. 395).

2.2 What does a GVA record contain?

For each incident, GVA records:

- Date of the incident
- Geocoded location (city, county, state, and coordinates)
- Available victim- and perpetrator-level information (e.g., name, age, sex)
- Incident type (e.g., "Shot – Wounded/Injured" or "Shot – Dead")
- URL links to online sources documenting the incident

The inclusion of source URLs is particularly important: it allows researchers to independently verify each case and to mine sources for additional information not captured in GVA's fields (Sierra-Arévalo & Nix, 2020, p. 1045).

2.3 Why use GVA?

GVA has several advantages over government data sources like the FBI's Law Enforcement Officers Killed and Assaulted (LEOKA) data and the National Incident-Based Reporting System (NIBRS):

- **Timeliness.** GVA is updated daily; LEOKA data are typically released 16–18 months after year's end (Sierra-Arévalo et al., 2023, p. 396).
- **Coverage of nonfatal assaults.** LEOKA and similar sources focus heavily on fatalities. GVA captures both fatal and nonfatal firearm assaults — critical because nonfatal assaults account for, on average, 83% of all firearm assaults on police (Sierra-Arévalo et al., 2023, p. 395).
- **Independent sourcing.** GVA draws on media and public sources rather than police-generated reports, sidestepping some well-known reporting biases in LEOKA and NIBRS (Sierra-Arévalo & Nix, 2020, p. 1052).
- **Verifiability.** Each record includes URL links enabling independent case-by-case verification.

2.4 Limitations of GVA

GVA is not a cure-all. Key limitations to keep in mind:

- It is unclear how complete GVA's coverage is, or whether there is meaningful state-level variation in the reporting of violence against police (Sierra-Arévalo & Nix, 2020, p. 1053).
- GVA's expansive definition of gun violence means that raw data require careful cleaning before use. In Sierra-Arévalo and Nix (2020), nearly 26% of the total cases provided by GVA were excluded based on the project's inclusion criteria (p. 1053).
- GVA does not reliably record certain incident-level details (e.g., body armor, wound location, firearm caliber) that would be useful for some research questions (Sierra-Arévalo et al., 2022, p. 5).

3 Case Selection Criteria

The raw GVA data include many incidents that fall outside the scope of this project. The following criteria define our **analytic sample** and must be applied consistently during cleaning and coding.

3.1 Who is included?

We include **active, sworn law enforcement officers** employed by a **local or state agency that responds to calls for service**. This includes:

- Officers at city/county police departments and sheriff's offices
- State police officers
- Special jurisdiction officers: transit police, school/university police, tribal police, wildlife or park police
- Part-time or reserve officers who, while on duty, perform the same function as a full-time local or state officer
- Local or state officers shot while cooperating with federal agencies in specialized task force operations

We **exclude**:

- Federal law enforcement officers (e.g., FBI, DEA, ATF, U.S. Marshals) — they do not engage in routine patrol or respond to community calls for service
- Parole and probation agents
- Court officers

(Sources: Sierra-Arévalo & Nix, 2020, p. 1045; Sierra-Arévalo et al., 2023, p. 396)

3.2 What incidents are included?

We include cases in which an included officer was **on duty** and had their **person or on-person equipment** (e.g., radio, ballistic vest, ballistic shield, duty belt) struck by a **bullet or bullet fragment** fired from a **firearm** controlled by a **person who was not an on-duty police officer**.

"On-person equipment" explicitly excludes vehicles — a bullet striking a patrol car does not qualify.

We also include cases in which officers were grazed by a bullet, shot in their ballistic vest or other protective equipment, or had secondary equipment (e.g., radio, duty belt, boots) struck by bullets, bullet fragments, or shrapnel. These are included because "they are the outcome of suspects firing rounds at officers where the difference between minimal and significant injury is exceedingly slim" (Sierra-Arévalo & Nix, 2020, footnote 8, p. 1056).

We **exclude**:

- Off-duty firearm injuries
- Injuries caused by means other than a firearm (e.g., shrapnel from an explosion, pellet guns)
- Self-inflicted firearm injuries, whether accidental (e.g., training accidents) or intentional (e.g., suicides or suicide attempts)
- "Blue-on-blue" incidents in which one on-duty officer accidentally shot another
- Cases in which a suspect pointed a firearm at an officer but did not fire
- Cases in which shots were fired but no officer was struck

(Sources: Sierra-Arévalo & Nix, 2020, pp. 1045–1046; Sierra-Arévalo et al., 2023, p. 396)

---

4 The Cleaning and Coding Process

4.1 Starting point: raw GVA data

GVA provides data directly to the research team upon request for a specified date range. The raw data include all GVA-recorded incidents involving law enforcement officers being shot — this is a much broader universe than our analytic sample. Do not assume that any case in the raw data belongs in the final dataset without verification.

4.2 Step 1: Case-by-case verification

Every case must be independently verified by following the URL sources listed in the GVA record. This is not optional — it is the core of the cleaning process.

Because URLs for online media reports are sometimes inactive, you may need to conduct an internet search using the incident date, location, and any available officer names to find alternative sources (Sierra-Arévalo & Nix, 2020, p. 1045).

What to verify:

- That the incident actually occurred as described
- That an on-duty, qualifying officer was struck by a bullet or bullet fragment (not merely shot at)
- That the shooter was not another officer
- That the officer was on duty
- That the officer is employed by a qualifying agency type

4.3 Step 2: Applying inclusion criteria

As you verify each case, apply the inclusion/exclusion criteria described in Section 2. Cases that do not meet all criteria are excluded from the analytic sample. The reason for exclusion should be recorded (see Appendix A of Sierra-Arévalo & Nix, 2020 for a frequency table of exclusion reasons used in prior work).

Note that cases may be excluded for multiple reasons simultaneously (e.g., a federal officer who accidentally shot themselves).

4.4 Step 3: Coding outcomes

The primary outcome variable is whether the firearm assault was **fatal** (officer died as a result) or **nonfatal**. This is coded as a dichotomous variable (1 = fatal, 0 = nonfatal).

In Sierra-Arévalo et al. (2022), roughly 16% of firearm assaults on officers resulted in a fatality across the 2014–2020 study period (p. 2).

4.5 Step 4: Flagging uncertain cases

Cases that are unclear or difficult to code should be **flagged** rather than decided unilaterally. Flagged cases are reviewed by the PI to arrive at a final coding decision (Sierra-Arévalo & Nix, 2020, p. 1045). When in doubt, flag it.

4.6 Conservative defaults

When evidence is ambiguous, the project errs toward **conservative estimates of nonfatal firearm injury**. Specifically:

- Cases where media sources describe an officer as "wounded," "injured," or "hurt" but do not specifically confirm a gunshot injury from a bullet, shot, bullet fragment, or shrapnel are **excluded** (Sierra-Arévalo & Nix, 2020, p. 1045).
- Cases where it is unclear whether an officer shot themselves, was shot by a suspect, or was shot by another officer are **excluded** (Sierra-Arévalo & Nix, 2020, p. 1045).

---

5 Ambiguous Cases and Known Edge Types

The GVA data regularly produce cases that do not fit neatly into the inclusion/exclusion criteria. Below are the main edge-case types encountered in prior work. For any new ambiguous case not covered here, flag it for PI review.

5.1 "Blue-on-blue" with partial suspect involvement

In cases involving a struggle over a firearm, we **include** cases where an officer was shot during a struggle over a firearm, even if it is unclear whether it was the suspect or the officer who pulled the trigger — as long as the firearm was being controlled by the suspect at the time.

However, if other officers also opened fire during the incident and it is not explicitly stated that the shot officer was hit by a round from the firearm over which the struggle occurred, the case is **excluded** on the grounds that a "blue-on-blue" / "friendly fire" incident cannot be ruled out (Sierra-Arévalo & Nix, 2020, footnote 9, p. 1056).

5.2 Simultaneous suspect and officer gunfire (the Helus case)

One of the clearest illustrations of edge-case complexity in this project: Sergeant Ron Helus of the Ventura County Sheriff's Department was shot six times while responding to an active shooter situation. Five shots were fired by the suspect; a coroner's report found that the sixth wound — to the heart, and ruled as the cause of death — was caused by a rifle round fired by another officer.

This case was **retained** as a nonfatal shooting by a suspect. The reasoning: even though one of the shots came from another officer, we could not conclusively determine whether the deputy would have survived had he not been shot through the heart by the fellow officer. Put differently, the case could not be unambiguously classified as a "blue-on-blue" death (Sierra-Arévalo & Nix, 2020, p. 1053).

This example underscores that some coding decisions will be imperfect and that "other researchers coding the same case might well have made a different coding decision" (Sierra-Arévalo & Nix, 2020, p. 1053). The goal is not perfection but consistency and transparency. All such decisions should be documented.

5.3 Unverifiable cases

Cases for which no online sources can be found to verify GVA-listed information are coded as **"unverifiable"** and excluded from the analytic sample (Sierra-Arévalo & Nix, 2020, footnote 11, p. 1056).

5.4 Unclear cases

Cases where available sources do not provide sufficient information to definitively determine who shot an officer are coded as **"unclear"** and excluded. Two examples from prior work:

- An officer nonfatally shot in the hand during a search warrant entry — sources indicate a suspect pointed a firearm at officers and was shot by officers, but do not confirm the suspect ever fired.
- An officer shot in the foot during a struggle with a suspect in a bar — a second officer present also fired, and the suspect's firearm discharged once; sources did not establish whose weapon struck the officer.

In both cases, the shot could plausibly have come from another officer, warranting exclusion (Sierra-Arévalo & Nix, 2020, footnote 11, p. 1056).

5.5 Cases not in GVA

Independent verification sometimes reveals incidents that meet our inclusion criteria but were **not** recorded in the GVA data. These cases should be added to the dataset and flagged as PI-identified cases not originally in GVA. In the 2020 paper, 15 such cases were found during independent verification, 11 of which were retained (Sierra-Arévalo & Nix, 2020, footnote 11, p. 1056).

---

6 A Note on Data Quality and the Importance of Manual Verification

The broader significance of careful manual verification cannot be overstated. As Sierra-Arévalo and Nix (2020) note, two high-profile studies using media- and crowdsourced data on police shootings came under scrutiny in 2019 for coding errors and debatable decisions — errors that, once corrected, reduced reported effects to statistical nonsignificance (p. 1054).

Manual, case-by-case verification is the standard this project holds itself to. It is time-consuming but essential. The richness of GVA data — particularly the source URLs attached to each case — makes this kind of verification possible and is one of GVA's key advantages over other data sources.

---

7 Key Papers

The following publications from this project provide the most detailed documentation of the data source, case selection, and coding process. When in doubt, consult the methods sections of these papers directly.

- **Sierra-Arévalo, M., & Nix, J. (2020).** Gun victimization in the line of duty: Fatal and nonfatal firearm assaults on police officers in the United States, 2014–2019. *Criminology & Public Policy, 19*(3), 1041–1066. — The foundational paper; contains the most detailed description of the GVA data source, case selection criteria, and coding decisions (including the edge-case discussion).
- **Sierra-Arévalo, M., Nix, J., & O'Guinn, B. (2022).** A national analysis of trauma care proximity and firearm assault survival among U.S. police. *Police Practice and Research.* — Uses the same case selection strategy as the 2020 paper, extended through 2020.

- **Sierra-Arévalo, M., Nix, J., & Mourtgos, S.M. (2023).** The "war on cops," retaliatory violence, and the murder of George Floyd. *Criminology, 61*(3). — Extends the dataset through 2020 and uses daily-level data; provides additional context on GVA's advantages relative to LEOKA and NIBRS.

---

8 Column-by-Column Coding Guidance

This section provides detailed instructions for coding each column in the project spreadsheet. The spreadsheet contains both raw GVA fields (populated by GVA prior to delivery) and project-added fields (coded by the research team). Coders are responsible only for the project-added fields described below. Raw GVA fields should not be altered.

8.1 General workflow

Begin each case by locating the `Sources` column, which contains the URLs that GVA recorded for that incident. These URLs may not be formatted as clickable hyperlinks within the spreadsheet. A practical workaround is to copy the entire cell contents into a Word document, where the URLs will become active and can be clicked individually.

Read each source carefully before coding any fields. The goal is to corroborate the information already present in the raw GVA fields and to gather additional details needed for the project-added columns. This process of cross-checking coded information against what can be independently found online applies to every field. When sources conflict with one another on key facts, note the discrepancy explicitly in the `Notes` field (see below).

When URLs are dead or inactive, do not simply skip the case. Conduct an internet search using the incident date, city/state, and any officer names listed in the raw data to locate alternative sources. When you find a corroborating source that is not listed in the original GVA data, record its URL in the `Notes` field along with a brief explanation of what information it provided.

When in doubt about any coding decision, flag the case for PI review rather than deciding unilaterally. A flag can be as simple as a note in the `Notes` field saying "unclear — flagged for PI review." Consistency and transparency are more important than speed.

8.2 `Status2`

Type: Categorical (controlled vocabulary)
Valid values: `Fatal`, `Non-fatal`

This is a cleaned re-coding of GVA's raw `Status` field. Code as follows:

- `Fatal` — the officer died as a result of gunshot wounds sustained in the line of duty
- `Non-fatal` — the officer was wounded but survived

`Status2` should reflect the **ultimate outcome** of the shooting. If an officer was initially reported as injured but later died from their wounds, the correct code is `Fatal`.

8.3 notactiveswornlocalstate

Type: Binary flag
Valid values: 0, 1

Code 1 if the victim is **not** an active, sworn, local or state law enforcement officer. This is a catch-all flag for any officer-type violation, including:

- Federal law enforcement officers (FBI, DEA, ATF, U.S. Marshals, Park Police, etc.)
- Retired officers — treated as civilians, not as off-duty officers. Do **not** code retired officers as offduty = 1; use notactiveswornlocalstate = 1
- Corrections officers
- Court officers
- Security guards or private security personnel mistakenly included in the GVA data

This column captures officer *type and status*, not whether they were on or off duty — that is handled separately by offduty.

Effect on sample: notactiveswornlocalstate = 1 → ToRemove = 1

8.4 rank

Type: Open text

Record the rank of the officer as reported in the source articles. Common values: Officer, Deputy, Sergeant, Detective, Trooper, Corporal, Lieutenant, Captain, Sheriff, Police Chief, Marshal.

Apply common-sense consistency checks using agency context:

- **Sheriff's office** employees are typically Deputy or a ranked variant — rarely just "Officer"
- Trooper is almost exclusively used by **state policing agencies**
- **Local police department** officers are typically Officer or a higher rank

When rank is not specified in any available source, use N/A.

8.5 agencytype

Type: Categorical (controlled vocabulary)
Valid values: Local, Sheriff, State, Special (X), Federal, Corrections, Unknown

Record the type of agency for which the officer works. For special jurisdiction agencies, use Special followed by the specific type in parentheses — e.g., Special (University), Special (School), Special (Park), Special (Housing Authority), Special (Transit), Special (Tribal). Be as specific as possible; look for examples in prior rows of the dataset when uncertain.

Federal and Corrections will typically co-occur with notactiveswornlocalstate = 1 and ToRemove = 1. Use Unknown only when the agency type genuinely cannot be determined from any available source.

8.6 `agencyname`

Type: Open text

Record the full official name of the agency. Use the agency's own official name as it appears on their website or official social media, cross-checking against incident sources. When sources conflict (e.g., "Sheriff's Office" vs. "Sheriff's Department"), defer to the agency's own official usage.

Common naming conventions:

- Local departments: `[City] Police Department`
- Sheriff's offices: `[County] County Sheriff's Office` or `[County] County Sheriff's Department`
- State agencies: `[State] State Police` or `[State] Highway Patrol`

When the name cannot be confirmed, use `Unknown` and flag in `Notes`.

8.7 `offduty`

Type: Binary flag
Valid values: `0`, `1`

Code `1` if the officer was shot while **not on the clock** in their law enforcement role, regardless of what they were doing — whether passively victimized or actively intervening in a criminal incident.

Do not confuse off-duty officers with retired officers:

- Currently employed officer, shot off duty → `offduty = 1`, `notactiveswornlocalstate = 0`
- Retired officer who intervenes → `notactiveswornlocalstate = 1`, `offduty = 0`

Effect on sample: `offduty = 1` → `ToRemove = 1`

8.8 `training`

Type: Binary flag
Valid values: `0`, `1`

Code `1` if the officer was shot during a **training exercise** (e.g., academy training, firearms qualification, simulated scenario). These officers are technically on the clock, but the shooting occurred outside normal operational duties and is treated as a distinct category.

Effect on sample: `training = 1` → `ToRemove = 1`

8.9 `blueonblue`

Type: Binary flag
Valid values: `0`, `1`

Code `1` if the officer was shot by **another officer**. Most commonly occurs in on-duty crossfire situations. In rare cases, `offduty = 1` and `blueonblue = 1` may both apply (e.g., an off-duty officer shot by an on-duty officer).

Note: `blueonblue` is a binary flag on the victim record; `type_new` captures the nature of the shooting from the incident perspective.

Effect on sample: `blueonblue = 1` → `ToRemove = 1`

8.10 `Error`

Type: Binary flag
Valid values: `0`, `1`

Code `1` if there is a **fundamental error in the underlying GVA raw data** — meaning the case should not have been included in the GVA pull at all. This is distinct from cases that are disqualifying under our inclusion criteria. Reserved for full GVA misclassifications, such as:

- A person listed as an officer who was actually a suspect
- A security guard listed as a police officer
- A person listed as having been shot who was in fact the shooter
- A case where sources reveal no evidence that any officer was struck

In most cases `Error = 1` also means `ToRemove = 1`, but not always — e.g., a PI-identified case added during independent research may receive `Error = 1` with an explanatory `ErrorDetail` but `ToRemove = 0` if the case meets all inclusion criteria.

When `Error = 0`, leave `ErrorDetail` as `N/A`.

8.11 `ErrorDetail`

Type: Open text

When `Error = 1`, provide a concise plain-text description of the specific error. A reviewer should be able to understand the nature of the error without consulting the sources. Leave as `N/A` when `Error = 0`.

8.12 `ToRemove`

Type: Binary flag
Valid values: `0`, `1`

Code `1` if this case should be **excluded from the analytic sample**. Triggered by any of the following:

- notactiveswornlocalstate = 1
- offduty = 1
- training = 1
- blueonblue = 1
- type_new is anything other than Suspect-inflicted or Suspect-inflicted; Accidental
- Error = 1 in most cases (read ErrorDetail to confirm)

Code all applicable flags accurately — ToRemove is the summary removal flag, not a substitute for completing the individual columns.

Important — duplicates: Apply ToRemove = 1 to the duplicate row only, not to the original. Verify which row is the original before assigning ToRemove.

8.13 type_new

Type: Categorical (controlled vocabulary)

Records the nature of the shooting. Use values **exactly as written** — capitalization and punctuation matter for downstream cleaning.

Value	Definition
Suspect-inflicted	Officer intentionally shot by a suspect. Most common type.
Suspect-inflicted; Accidental	Suspect shot the officer unintentionally — e.g., misidentified officer as a threat and fired through a door.
Suspect-inflicted; Blue-On-Blue	Shooter was a police officer acting as a suspect/aggressor. Exceedingly rare.
Accidental; Blue-On-Blue	Officer accidentally shot by another on-duty officer. Most common blue-on-blue type.
Accidental; Self-inflicted	Officer accidentally shot themselves. Often co-occurs with training = 1.
Self-inflicted; Suicide	Officer intentionally shot themselves.
N/A	Type cannot be determined from available sources. Typically co-occurs with Error = 1 and ToRemove = 1.

Struggle over a firearm: Code as Suspect-inflicted if the firearm was controlled by the suspect at the time of discharge, even if it is unclear who pulled the trigger. However, if other officers also opened fire and it cannot be confirmed the shot officer was struck by the suspect's weapon specifically, code N/A and flag for PI review.

Effect on sample: Suspect-inflicted and Suspect-inflicted; Accidental do **not** trigger ToRemove = 1. All other values do.

8.14 Notes

Type: Open text (semi-structured)

The Notes field serves as both a structured case summary and a documentation trail for sourcing decisions.

Standard format:

[Agency type] [Rank] [fatally/non-fatally] shot [wound location if known] [incident context].

Examples:

> *"Local Officer non-fatally shot in the leg while responding to a domestic disturbance call."*
>
> *"State Trooper fatally shot during a traffic stop."*
>
> *"Sheriff's Deputy non-fatally shot in the chest (stopped by ballistic vest) and in the left shoulder while serving an arrest warrant."*

Render agency type and rank consistently with `agencytype` and `rank`:

- `Local` → "Local Officer / Local Sergeant" etc.
- `Sheriff` → "Sheriff's Deputy / Sheriff's Sergeant" etc.
- `State` → "State Trooper"
- `Special (School)` → "School Police Officer / School Police Sergeant" etc.

Wound location: Record only what sources explicitly state. When a bullet is stopped by a vest, record both (e.g., "shot in the chest (stopped by ballistic vest) and in the left shoulder"). When equipment rather than person is struck, note that specifically. Do not infer.

Multi-officer incidents: Each record in a multi-officer incident must reference the other officers involved. A reader looking at any single record should be able to reconstruct the full picture of the incident from that note alone. Notes across records for the same incident should be mutually consistent.

Example (two-officer incident):

> Record 1: *"Sheriff's Deputy non-fatally shot in the arm and back after responding to a domestic disturbance call. Another Sheriff's Deputy was non-fatally shot in the leg."*
>
> Record 2: *"Sheriff's Deputy non-fatally shot in the leg after responding to a domestic disturbance call. Another Sheriff's Deputy was non-fatally shot in the arm and back."*

Source documentation: After the case summary, record any corroborating URLs found during independent research that are not in the original GVA `Sources` field, with brief context explaining what each URL provided. Also note when a GVA-listed URL does not correspond to the incident.

Discrepancy flagging: When sources conflict on key facts, explain the discrepancy and provide relevant URLs so a reviewer can assess the evidence independently.

8.15 `Duplicate`

Type: Binary flag
Valid values: `0`, `1`

Code `1` if this row is a duplicate of another row already in the dataset. When `Duplicate = 1`, also complete `duplicate_of`. Apply `ToRemove = 1` to the duplicate row only — not to the original.

8.16 duplicate_of

Type: Numeric (Incident ID)

When `Duplicate = 1`, record the Incident ID of the original case this row duplicates. Leave `N/A` when `Duplicate = 0`.

8.17 record_added

Type: Binary flag
Valid values: `0`, `1`

Code `1` if this row was not present in the original GVA data pull and was added during the cleaning process. Two scenarios:

1. **Additional officer(s) from an existing incident:** GVA recorded one officer but independent research revealed additional officers were also shot. Each additional officer gets a new row with `record_added = 1`. These rows **inherit the existing Incident ID** — no new ID is needed.
2. **Wholly new incident:** An incident not recorded anywhere in the raw GVA data. All fields must be completed as if coding a normal GVA record, including coordinates. This is the one `record_added` scenario that **requires a new Incident ID** (see `incident_id_changed` below).

8.18 incident_id_changed

Type: Binary flag
Valid values: `0`, `1`

Code `1` if the Incident ID was changed from what GVA originally assigned. Most commonly necessary when two cases were grouped under the same ID but represent separate incidents. Record the new ID in `Incident ID` and preserve the original in `original_id`.

When to create a new Incident ID: Only in two situations — (1) a genuinely missing incident found during independent research, or (2) a case incorrectly grouped under an existing ID that is a separate incident not represented anywhere else in the dataset.

New ID convention (implemented for 2024–2025 data coding beginning February 2026; prior datasets used a different ad hoc approach):

> **Format:** `999YYYYMMDDX`
>
> - `999` — fixed prefix distinguishing synthetic IDs from GVA-assigned IDs
> - `YYYYMMDD` — incident date
> - `X` — trailing digit (1, 2, 3…) for multiple added cases on the same date
>
> *Example: Two added cases on March 12, 2015 → `999201503121` and `999201503122`*

This produces a 12-digit integer outside GVA's current ID range, eliminating collision risk. The `999` prefix makes synthetic IDs immediately recognizable.

For `incident_id_changed` cases, flag and defer final ID assignment to the PI if there is any uncertainty about whether the split is warranted.

8.19 `original_id`

Type: Numeric (Incident ID)

When `incident_id_changed = 1`, record the original GVA-assigned Incident ID here. Leave `N/A` when `incident_id_changed = 0`.

---

9 Quick Reference: `ToRemove` Triggers

Condition	Column
Not an active, sworn, local/state officer	`notactiveswornlocalstate = 1`
Officer was off duty	`offduty = 1`
Shooting occurred during training	`training = 1`
Officer was shot by another officer	`blueonblue = 1`
Shooting type is not suspect-inflicted (intentional or accidental)	`type_new` ≠ `Suspect-inflicted` or `Suspect-inflicted; Accidental`
Fundamental GVA data error (in most cases)	`Error = 1` — confirm via `ErrorDetail`

A case can trigger multiple conditions simultaneously. Code all applicable flags accurately — `ToRemove` is the summary removal flag, not a substitute for completing the individual columns.

Document prepared for internal project use. Direct questions to the PI.